目　　录

P-0-1 矩阵、矩阵的和及积

有限个标以顺序的元素[1] 排列成的矩形称为矩阵 (matrix). 元素的横排称为行(row), 纵排称为列(column). 行从上数分别称为第1行、第2行……, 第 m 行; 列从左数分别称为第1列、第2列, ……, 第 n 列. 位于第 i 行第 j 列的元素称为(i, j)元素, 把 i, j 写为下角可表示为 a_{ij}. 因此,典型的矩阵可以表示为

$$\boldsymbol{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \tag{1}$$

为了简便起见,矩阵常用(i, j)元素代表,有时也写成

$$\boldsymbol{A}=[a_{ij}] \tag{2}$$

具有 m 行 n 列的矩阵, 称为 m 行 n 列矩阵,或简称为 $m\times n$ 矩阵, 有时写成 $\boldsymbol{A}=m\times n$, 等等. 特别是,把 $\boldsymbol{A}=n\times n$, 即行数和列数相等的矩阵称作方阵 (Square matrix).

在方阵中,(i, i)元素如$(i=1, 2, \cdots, n)$称为主对角线元素.

元素全为0的 $m\times n$ 矩阵称作 $m\times n$ 零阵,写成 $\mathbf{0}_{m\times n}$, 或简写成 $\mathbf{0}$.

仅主对角线元素为1, 而其余元素全为0的 $n\times n$ 方阵, 称为 n 阶单位阵,写成 $\boldsymbol{I}_n$, 或简写成 $\boldsymbol{I}$.

矩阵的和及积

以 $\boldsymbol{A}=[a_{ij}]=m\times n$, $\boldsymbol{B}=[b_{ij}]=m\times n$ 两个矩阵对应元素的和为元素而构成的矩阵, 定义为 $\boldsymbol{A}$、$\boldsymbol{B}$ 的和, 即[2]

$$\boldsymbol{A}+\boldsymbol{B}\triangleq[a_{ij}+b_{ij}]=m\times n \tag{3}$$

设 $\boldsymbol{A}=[a_{ij}]=m\times n$ 和 $\boldsymbol{B}=[b_{ij}]=p\times q$, 当 $\boldsymbol{A}$ 的列数等于 $\boldsymbol{B}$ 的行数(即 $n=p$)时, 其积 $\boldsymbol{AB}$ 可定义如下

$$\boldsymbol{AB}\triangleq\left[\sum_{k=1}^{n} a_{ik}b_{kj}\right]=m\times q \tag{4}$$

对于矩阵的积必须注意: (i)在 $n=p$, $m=q$ 均成立的情况下, 可以同时定义乘积 $\boldsymbol{AB}$ 及与其顺序相反的乘积 $\boldsymbol{BA}$, $\boldsymbol{AB}=m\times m$, $\boldsymbol{BA}=n\times n$. 一般, $\boldsymbol{AB}$ 与 $\boldsymbol{BA}$ 是大小不同的矩阵. (ii)仅当 $m=n=p=q$, 即 $\boldsymbol{A}$, $\boldsymbol{B}$ 是大小相同的方阵时, $\boldsymbol{AB}$ 和 $\boldsymbol{BA}$ 才是大小相同的矩阵. (iii)即便在这种情况下,一般 $\boldsymbol{AB}=\boldsymbol{BA}$ 也不成立. (iv)尽管 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 都不是零阵,但是乘积 $\boldsymbol{AB}$ 有时变成零阵,等等.

数 α 和矩阵 $\boldsymbol{A}=[a_{ij}]=m\times n$ 之积可定义为

$$\alpha\boldsymbol{A}\triangleq[\alpha a_{ij}]=m\times n \tag{5}$$

---

1) 这里可以将元素理解成实数或复数.

2) $\triangleq$ 意味着根据定义相等.

A-0-1 自动控制和矩阵

大家知道，从1950年下半年到1960年初，自动控制这门科学(自动控制系统中问题的提出、系统的表示、分析和设计等等)得到很大发展,形成所谓的现代控制理论. 从它的产生到现在,经历了十年以上的各种考验，显然今后它已是一种肯定的科学，而且和所谓的经典控制理论的关系也逐渐清楚了.

因此可以说,若将现代控制理论和经典控制理论作为两门控制科学,从自动控制的课程设置来说,不掌握经典控制理论就无法学习现代控制理论. 根据这个观点,有人把两者结合起来构成一门控制科学的体系. 现在已经有一些按照这种观点写成的教科书[1~5],它们反映了这样一种倾向.

那末,为了掌握按照这种新的体系写成的控制理论，从数学基础来看，除了必须掌握经典控制理论的基础拉普拉斯变换以外，还要求具备线性代数、特别是矩阵理论的知识,这个道理越来越明显. 在现代控制理论中,一般都是考虑多变量控制,而不像经典控制理论那样主要是讨论单变量控制. 因此,利用矩阵写出多变量系统的方程式,不但运算时符号简便,而且可以利用矩阵的各种性质进行理论上的进一步研究(不利用矩阵是非常困难的). 用矩阵写成的方程式,除了极简单的情况而外,一般不便于手算,而具有适合于在计算机中处理的特点.

现代控制理论以所谓的动力学系统的理论为基础,它由线性多变量系统的系统设计、最优控制、状态估计、系统识别及稳定性分析等部分组成. 全面地介绍这些内容所涉及到的矩阵理论,作者无论如何也办不到，而且也不是写作本书的目的. 因此，这里我们仅选取表1中所列出的一些项目,暂且称其为自动控制中的基本理论,介绍作为其基础的矩阵理论[1)].

表 1

线性多变量系统	结构理论(可控性、可观测性、实现问题等等)
	控制问题(二次型性能指标的最优控制、模态控制、解耦控制、极点配置、状态反馈的稳定化等等)
	灵敏度分析
	状态估计(观测器、卡尔曼滤波)
	稳定性分析(李亚普诺夫稳定性理论、有界输入-有界输出稳定性)
	响应的数值分析(e^{At}的计算等等)
非线性系统的稳定性分析(李亚普诺夫稳定性理论、频域的稳定性判据、卢里叶问题等)	
线性顺序系统的分析(响应的表示、可控性、可观测性等等)	

1) 除了本书所介绍的内容以外,自动控制中所涉及到的非常重要的矩阵理论，还有数学规划(线性规划、平方规划、共轭梯度法、梯度投影法等等),读者可以参考文献[11~18].

虽然表中大部分内容在上述教科书中都有介绍，但是目前还没有一本合适的线性代数参考书，能系统地介绍作为其基础的矩阵理论，而且在自动控制教科书的附录中所能看到的关于矩阵的介绍，大多只限于最基本的内容．从这个角度来说，本书对读者若能有一些帮助，作者也感到安慰．

因为本书的写作对象是从事自动控制工作的工程技术人员，所以在讲述某些数学内容的同时，尽可能结合表1所列出的课题，介绍矩阵理论在自动控制中的应用．

B-0-1 排列和矩阵

我们所运用的数的范围，从自然数开始逐渐扩大成整数、有理数、实数、复数．其中，复数由实部和虚部两个元素组成．将其推广，把更多个(不限于两个)数(整数、有理数、实数、复数)在一定条件下的排列看成更广义的数．矩阵就是这种把有限个数按一定顺序排成m行n列的数，按着这种考虑可以将“数”推广．像这样，可以看成“数的排列 (array)”是矩阵的第一个特点．

但是，这里要注意两点．第一，一谈到数，人们就自然地会认为可以对它进行四则运算(加，减，乘，除)．把这种观点原封不动地拿到作为广义的“数”矩阵，并不完全合适．本书开始部分谈到矩阵的运算，例如在矩阵的乘积 $\boldsymbol{AB}$ 中可以看到，其顺序不一定可以交换，因此就有必要注意与运算有关的性质．因此还有一点，将矩阵 $\boldsymbol{A}=[a_{ij}]$ 看成数的排列，其元素 a_{ij} 不只限于实数及复数，而可以更自由地考虑，它可以是由实数、复数按照四则运算得到的，例如a_{ij} 可以是$(b_nx^n+b_{n-1}x^{n-1}+\cdots+b_0)/(a_nx^n+a_{n-1}x^{n-1}+\cdots+a_0)$这样一个有理数．象这样，可以将矩阵作更广泛的考虑，对于在自动控制中的应用也是很方便的．

为了对上述各点作稍为深入一步的讨论和加以整理，利用代数学中群、环、体的概念是很方便的[1)]．

[定义 1]　当集合 G 满足下列公理时，称 G 为群(group)．

a) 对于 G 的任意两个元 α, β, 可以定义某一个运算“$\circ$”, $\alpha\circ\beta$ 作为 G 中的元唯一确定．(自闭性，唯一性)

b) 对于 G 的任意三个元 α, β, γ, $(\alpha\circ\beta)\circ\gamma=\alpha\circ(\beta\circ\gamma)$成立．(结合律)

c) G 中存在元 e, 对于 G 的任意元 α, $e\circ\alpha=\alpha\circ e=\alpha$ 成立．(单位元)

d) 对于 G 的任意元 α, 存在满足 $\alpha^{-1}\circ\alpha=\alpha\circ\alpha^{-1}=e$ 的元 α^{-1}, 而且 α^{-1} 唯一确定．(逆元)

在上面定义中对于运算“$\circ$”没有假定交换律成立，当下列交换律

e) 对于 G 的任意两个元 α, β, $\alpha\circ\beta=\beta\circ\alpha$ 成立．

成立时，该群称为可换群[2)] (阿贝尔群)．

当运算“$\circ$”是加法时，满足上面 a)～e) 的群称为加群 (即关于加法的可换群)．在加群中，单位元 e 写成 0, 叫做零元，而且将 α^{-1} 写成 $-\alpha$．当运算“$\circ$”是乘法时，满足 a)～e) 的群称作关于乘法的可换群，这时单位元 e 可以写成 1．此外，通常将 $\alpha\circ\beta$ 写成 $\alpha\beta$．

---

1) 关于代数学，除了传统性的[6,7] 以外，K. L. 加多纳著的《近代代数学的发现》(ダイヤモント社出版)也是一本很好的入门参考书．

2) 使用逻辑符号 $\forall$ (对于所有的…)和表示所属关系的符号 $\in$ (属于…)，上述定义可以写成

$${}^\forall\alpha,\ \beta\in G，\text{可以定义 } \alpha\circ\beta$$

a) $\alpha\circ\beta\in G$

b) $(\alpha\circ\beta)\circ\gamma=\alpha\circ(\beta\circ\gamma)$, ${}^\forall\alpha,\ \beta,\ \gamma\in G$

c) G 中存在元 e, $e\circ\alpha=\alpha\circ e=\alpha$, ${}^\forall\alpha\in G$

d) 对于 ${}^\forall\alpha\in G$, 有满足 $\alpha^{-1}\circ\alpha=\alpha\circ\alpha^{-1}=e$ 的元 $\alpha^{-1}\in G$ 存在

e) $\alpha\circ\beta=\beta\circ\alpha$, ${}^\forall\alpha,\ \beta\in G$

[例 1]　实数的集合 R 是加群，R 中除掉 0 后的集合，变成关于乘法的可换群.

[定义 2]　当集合 $\widetilde{R}$ 满足下列公理时，称 $\widetilde{R}$ 为环(ring).

a) 对于 $\widetilde{R}$ 的任意两个元 α,β，可以定义加法＋和乘法·，而且 $\alpha+\beta$、$\alpha\beta$ 皆唯一确定，

$$\alpha+\beta\in\widetilde{R},\ \ \alpha\beta\in\widetilde{R},\ \ {}^{\forall}\alpha,\ \beta\in\widetilde{R}$$

b) $\widetilde{R}$ 是加群.

c) 对于乘法·，结合律成立.

$$(\alpha\beta)\gamma=\alpha(\beta\gamma),\ \ {}^{\forall}\alpha、\beta,\ \gamma\in\widetilde{R}$$

d) 对于加法和乘法都满足分配律

$$\alpha(\beta+\gamma)=\alpha\beta+\alpha\gamma\quad {}^{\forall}\alpha,\ \beta,\ \gamma\in\widetilde{R}$$

$$(\alpha+\beta)\gamma=\alpha\gamma+\beta\gamma\quad {}^{\forall}\alpha,\ \beta,\ \gamma\in\widetilde{R}$$

在上述环的定义中，根据条件 b)，$\alpha+\beta=\beta+\alpha$. 但不要求关于乘法的可换性. 作为附加条件，把满足乘法交换律的环叫做可换环.

[例 2]　全体整数的集合 $Z\{0,\ \pm1,\ \pm2,\ \cdots\}$，是关于普通加法、乘法的可换环. 偶数的集合 $Z_e\{0,\ \pm2,\ \pm4,\ \cdots\}$ 也是同样.

[例 3]　设 $\widetilde{R}$ 为任意环，将 $\widetilde{R}$ 中的元排列成 n 行 n 列方阵

$$\boldsymbol{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

等，$\boldsymbol{A}$ 称为环 $\widetilde{R}$ 上的矩阵. 当 n 固定时，这样的矩阵全体的集合，我们用 $\widetilde{R}^{n\times n}$ 表示. 在 $\widetilde{R}^{n\times n}$ 中应用 P-0-1 中(3)，(4)式定义的加法、乘法，可以证明 $\widetilde{R}^{n\times n}$ 也是环(试证明)，也叫做矩阵环. 当然，既使 $\widetilde{R}$ 是可换环，一般 $\widetilde{R}^{n\times n}$ 也不一定是可换环.

[定义 3]　当集合 F 满足下列公理时，称 F 为体(field).

a) F 是可换环.

b) F 中除 0 以外的元，对乘法构成可换群.

虽然体和环、群有关，可如上定义，但因体是今后经常用到的重要的概念，现再将其条件反复列举如下.

[定义 3′]　以 $\alpha,\ \beta,\ \gamma\cdots$ 为元的集合 F，当满足下列公理时，称 F 为体.

加法：
- a) 对于 F 的任意两个元 $\alpha,\ \beta,\ \alpha$ 和 β 的和可以唯一定义，$\alpha+\beta\in F$
- b) $\alpha+\beta=\beta+\alpha,\ {}^{\forall}\alpha,\ \beta\in F$(加法的交换律)
- c) $\alpha+(\beta+\gamma)=(\alpha+\beta)+\gamma,\ {}^{\forall}\alpha,\ \beta,\ \gamma\in F$(加法的结合律)
- d) F中存在 0 元，$\alpha+0=\alpha,\ {}^{\forall}\alpha\in F$ 成立.
- e) 对于任意的 $\alpha\in F$，在 F 中存在 $-\alpha$ 元，$\alpha+(-\alpha)=0$成立.

乘法：
- f) 对于 F 中的任意两个元 $\alpha,\ \beta,\ \alpha$ 和 β 的积 $\alpha\beta$ 可唯一定义，$\alpha\beta\in F$
- g) $\alpha\beta=\beta\alpha,\ {}^{\forall}\alpha,\ \beta\in F$(乘法的交换律)
- h) $\alpha(\beta\gamma)=(\alpha\beta)\gamma,\ {}^{\forall}\alpha,\ \beta,\ \gamma\in F$(乘法的结合律)
- i) F 中存在单位元 1，$\alpha 1=\alpha,\ {}^{\forall}\alpha\in F$ 成立.
- j) 对于 F 中的任意 $\alpha\neq0$ 的元，在 F 中有用 α^{-1} 表示的逆元存在，$\alpha\alpha^{-1}=\alpha^{-1}\alpha=1$ 成立.

对于加法和乘法还满足下列分配律

k) $\alpha(\beta+\gamma)=\alpha\beta+\alpha\gamma$, $^\forall\alpha, \beta, \gamma\in F$(分配律)

[例 4] 全体有理数的集合，对于普通加法、乘法构成体.

[例 5] 全体实数的集合 R, 复数的集合 C 均构成体.

(问题 1) 试证明体 F 中加法的单位元 0 及乘法的单位元 1 均唯一确定.

由定义 3′ 可见，体就是可以任意地对其元做有限次加减乘除的集合. 只知道自然数的儿童不会做由 2 减去 5, 用 3 除尽 2 而感到不自由，这是因为自然数或整数不能构成体. 另方面，在有理数、实数、复数等体上，我们随心所欲地用惯了四则运算，而对于矩阵的运算也感到不自由(不满足乘法的交换律，逆矩阵不一定存在，等等)，如例 3 所示，这是因为矩阵最多只能构成环.

(问题 2) 试证明例 3 中的方阵改成 m 行 n 列的长方矩阵$(m\neq n)$时又如何?

(问题 3) 在例 3 中将环 $\tilde{R}$ 换成体 F, 将 F 的元排列成 n 行 n 列的方阵，这些矩阵的全体用 $F^{n\times n}$ 表示，试证明 $F^{n\times n}$ 关于矩阵的加法、乘法能否构成体.

群、环、体的概念不仅与上述的数有关，而且有更为广泛的应用. 下面结合经常用到的重要的部分举几个例子.

[例 6] 设 F 为任意体，由该体的元 a_k 和变量 s 组成的多项式

$$a_0+a_1s+\cdots+a_ns^n=\sum_{i=0}^{n}a_is^i \tag{1}$$

的全体，用 $P(s, F)$ 表示.

对于两个多项式

$$f(s)=\sum_{i=0}^{n}a_is^i,\quad g(s)=\sum_{i=0}^{n}b_is^i,$$

在定义其和

$$f(s)+g(s)\triangleq\sum_{i=0}^{n}(a_i+b_i)s^i \tag{2}$$

及积

$$f(s)g(s)\triangleq\sum_{i=0}^{2n}c_is^i,\quad c_i\triangleq\sum_{\substack{j+k=i\\ j,k\leq n}}a_jb_i \tag{3}$$

时，$P(s, F)$便构成环.

[例 7] 设 F 为任意体，由 F 的元 a_k, b_k 和变量 s 构成有理式.

$$\frac{b_0+b_1s+\cdots+b_ns^m}{a_0+a_1s+\cdots+a_ns^n}=\frac{n(s)}{d(s)} \tag{4}$$

的全体，用 $Q(s, F)$表示，当定义两个有理函数的和及积

$$\frac{n_1(s)}{d_1(s)}+\frac{n_2(s)}{d_2(s)}\triangleq\frac{n_1(s)d_2(s)+n_2(s)d_1(s)}{d_1(s)d_2(s)} \tag{5}$$

$$\frac{n_1(s)}{d_1(s)}\cdot\frac{n_2(s)}{d_2(s)}\triangleq\frac{n_1(s)n_2(s)}{d_1(s)d_2(s)} \tag{6}$$

时，$Q(s, F)$构成体.

[例 8] 设 S 为任意集合，$\tilde{R}$ 为任意环，并设定义在 S 上而在 $\tilde{R}$ 上取值的函数的全体为 Φ. 对于 $\varphi_1(t)$, $\varphi_2(t)\in\Phi$, 当定义其和及积

$$(\varphi_1+\varphi_2)(t)\triangleq\varphi_1(t)+\varphi_2(t) \tag{7}$$

$$(\varphi_1\varphi_2)(t) \triangleq \varphi_1(t)\varphi_2(t) \tag{8}$$

时，Φ 构成环．

在实轴上定义的连续(分段连续)实函数的全体，对于上述运算也构成环．

到此为止，以上所讨论的集合均具有无限个元．具有无限个元的体叫做无限体(infinite field)．与此相应，仅由有限个元构成的体叫做有限体(finite field)．下面举一些大家最熟悉的例子．

［例 9］ 在集合 $\{0, 1\}$ 上，将加法和乘法定义为：$0+0=0$, $0+1=1+0=1$, $1+1=0$, $0\cdot0=0\cdot1=1\cdot0=0$, $1\cdot1=1$, 则 $\{0, 1\}$ 构成体．

设 p 为质数，有限集合 $\{0, 1, \cdots, p-1\}$ 一般用 $GF(p)$ 表示．在 $GF(p)$ 上定义 mod p 运算(即当 $\alpha-\beta$ 是 p 的倍数时，认为 $\alpha=\beta$)时，$GF(p)$ 构成体，该体称为伽罗瓦体(Galois field)．

［例 10］ 在 $GF(3)$ 上的 mod 3 加法和乘法如下

加 法

	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

乘 法

	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

由此可见，$GF(3)$ 构成体．

(问题 4) 在 $GF(p)$ 上定义 mod p 运算，当 p 不是质数时，试以 $p=4$ 验证 $GF(p)$ 不是体．

容易看到，环和体的重要区别是在乘法运算方面．一般来说，环除了不满足乘法的交换律以外，乘法的单位元也不一定存在[1)](例如例 2 中的 Z_e)．既使假定单位元存在，对于任意 $\alpha\neq0$ 的元，使 $\alpha^{-1}\alpha=\alpha\alpha^{-1}=1$ 成立的 α^{-1} 也不一定定义在该环上[2)]．在例 2 的 Z 上，若 $\alpha=+1, -1$, 则 α^{-1} 分别为 $+1, -1$, 是环 Z 上的元，但是对于 Z 上的其它元，例如 $\alpha=2$ 时，α^{-1} 为 $1/2$, 它不是 Z 上的元．环 Z 上的 $+1, -1$, 其 α^{-1} 是 Z 上的元，这种元叫做可逆元．在环 $P(s, F)$ 上(例 6)，除了零次多项式(即体 F 的 0)以外，其余的元均为可逆元．

根据例 3, 将环 $\tilde{R}$ 中的元排列成 n 行 n 列方阵，该方阵的全体 $\tilde{R}^{n\times n}$ 也构成环．下面简单讲一下该环的性质．如果 $\tilde{R}$ 中存在单位元 1, 则 $\tilde{R}^{n\times n}$ 中也存在类似的单位元，即存在主对角线元素均为 1, 而其它元素均为 0 的 n 阶方阵．环 $\tilde{R}^{n\times n}$ 中*一定存在乘法的单位元．在这里考虑 $\tilde{R}$ 为可换环，而且为了简单起见，令 $n=2$, 我们来讨论一下 $\tilde{R}^{2\times2}$．对于 $\boldsymbol{A}\triangleq[a_{ij}]\in\tilde{R}^{2\times2}$, 若定义

$$\tilde{A}\triangleq\begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}\in\tilde{R}^{2\times2}$$

$$\Delta\triangleq a_{11}a_{22}-a_{21}a_{12}\in\tilde{R},$$

1) 有理数、实数、复数构成体是显然的，关于有理函数的集合 $Q(s, F)$ 及有限集合 $GF(p)$ 构成体的证明，在文献[10]中讲得很详细．

2) 若存在则唯一确定．

* 原文误为环 $F^{n\times n}$ 中．——译者注

则容易看出，它满足 $A\tilde{A}=\tilde{A}A=\Delta I$. 当且仅当 Δ 是 $\tilde{R}$ 的可逆元时，$\Delta^{-1}\in\tilde{R}$ 存在. A^{-1} 由

$$A^{-1}=\Delta^{-1}\tilde{A}\in\tilde{R}^{2\times 2}$$

给出，即 A 为 $\tilde{R}^{2\times 2}$ 的可逆元的充分必要条件是，Δ 是 $\tilde{R}$ 的可逆元.

一般，若矩阵 $A\in\tilde{R}^{n\times n}$ 为可逆元，则 A 称为幺模阵.

作为特殊情况，考虑体 F 上的 2×2 方阵，即令 $A\in F^{2\times 2}$. 因 $\tilde{A}\in F^{2\times 2}$, $\Delta\in F$, 若 $\Delta\neq 0$, 则 $\Delta^{-1}\in F$. 因而得

$$A^{-1}\triangleq\Delta^{-1}\tilde{A}\in F^{2\times 2},$$

即体 F 上的矩阵为可逆的充分必要条件是 $\Delta\neq 0$. 这种矩阵称为正则矩阵.

现举个例子，对于下列以整数为元素的矩阵

$$A=\begin{bmatrix}2 & 0\\ 0 & 1\end{bmatrix},$$

$\Delta=2$. 因此，将 A 看作 $R^{2\times 2}$ 上的矩阵时，是正则矩阵. 但因 $\Delta=2$ 不是 Z 上的可逆元，若将 A 看作 $Z^{2\times 2}$ 上的元，则不是幺模阵，实际上

$$A^{-1}=\begin{bmatrix}\frac{1}{2} & 0\\ 0 & 1\end{bmatrix}\in Z^{2\times 2}.$$

其次，考虑

$$A=\begin{bmatrix}2 & 1\\ 1 & 1\end{bmatrix},$$

在这种情况下，$\Delta=1$, 当然 $A\in R^{2\times 2}$ 是正则的. 因 $\Delta=1$ 是 Z 的可逆元，则 A 也是幺模阵，实际上

$$A^{-1}=\begin{bmatrix}1 & -1\\ -1 & 2\end{bmatrix}\in Z^{2\times 2}$$

成立.

另外，对于

$$A=\begin{bmatrix}s+1 & 0\\ 0 & s+2\end{bmatrix},$$

$\Delta=(s+1)(s+2)\neq 0$. 因此，将 A 看成 $Q(s, R)$ 上的矩阵时，是正则的. 但是将 A 看成 $P(s, R)$ 上的矩阵时，因 $(s+1)(s+2)$ 不是 $P(s, R)$ 的可逆元，则 A 不是幺模阵，实际上

$$A^{-1}=\begin{bmatrix}\frac{1}{s+1} & 0\\ 0 & \frac{1}{s+2}\end{bmatrix}\notin[P(s, R)]^{2\times 2}.$$

最后，对于

$$A=\begin{bmatrix}s+1 & s\\ s+3 & s+2\end{bmatrix},$$

$\Delta=2$, 因它是 $P(s, R)$ 上的可逆元，则 A 为幺模阵. 实际上

$$A^{-1}=\begin{bmatrix} \frac{1}{2}s+1 & -\frac{1}{2}s \\ -\frac{1}{2}s-\frac{3}{2} & \frac{1}{2}s+\frac{1}{2} \end{bmatrix}\in[P(s,\ R)]^{2\times2}$$

成立.

以上，抽象的内容讲了很多，现在大致整理一下自动控制中用到的矩阵和上述内容的关系．首先考虑大家所熟悉的线性时变系统的状态方程式和输出方程式

$$\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}+\boldsymbol{B}(t)\boldsymbol{u},\quad \boldsymbol{y}=\boldsymbol{C}(t)\boldsymbol{x}. \tag{9}$$

式中 $\boldsymbol{A}(t)$, $\boldsymbol{B}(t)$, $\boldsymbol{C}(t)$ 的元素都是定义在实轴上的函数．若状态变量是具有一定物理意义的变量，则亦为实函数(在例 8 中相当于 $S\in R^*$, $\tilde{R}=R$ 的情况)．其次，将所讨论的对象限于常系数线性系统

$$\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{u},\quad \boldsymbol{y}=\boldsymbol{C}\boldsymbol{x} \tag{10}$$

式中 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ 是以实数为元素的矩阵，即是实数体上的矩阵．其次，设起始条件 $\boldsymbol{x}(0)=\boldsymbol{0}$, 在 (10) 式中进行拉氏变换，求出输出与输入之比，便得到大家熟悉的传递函数矩阵 $\boldsymbol{G}(s)$

$$\boldsymbol{G}(s)=\boldsymbol{C}[s\boldsymbol{I}-\boldsymbol{A}]^{-1}\boldsymbol{B} \tag{11}$$

显然，式中 $\boldsymbol{G}(s)$ 是实系数有理函数体(例 7)上的矩阵，即 $\boldsymbol{G}(s)$ 的元素是 $Q(s,\ R)$ 的元．例如，$\boldsymbol{G}(s)$ 是具有下列形式

$$\boldsymbol{G}(s)=\begin{bmatrix} \dfrac{2}{s+1} & \dfrac{8}{s+2} \\ \dfrac{5}{(s+1)(s+3)} & \dfrac{1}{s+3} \end{bmatrix}$$

的矩阵．设该式右边分母的最小公倍多项式为 $q(s)$，利用多项式环(例 6) 上的矩阵 $\boldsymbol{P}(s)$ 可以写成

$$\boldsymbol{G}(s)=\frac{\boldsymbol{P}(s)}{q(s)}. \tag{12}$$

在上例中

$$q(s)=(s+1)(s+2)(s+3),$$

$$\boldsymbol{P}(s)=\begin{bmatrix} 2(s+2)(s+3) & 8(s+1)(s+3) \\ 5(s+3) & (s+1)(s+2) \end{bmatrix}.$$

有限体 $GF(p)$ (例 8, 例 9) 上的矩阵在自动控制中也获得应用．首先，它在系统识别时广泛采用的 M 系列信号的产生中就已用到．其次，在处理用作程序控制模型的线性顺序系统中，它也是不可缺少的．

如上所述，因为自动控制中要用到各种各样的矩阵，所以根据环及体的概念对它们进行整理具有重要意义．也就是说，基于体(或环)的共同性质得出的结论，虽然是由某个体(环)推导出来的，必然适合于其它所有的体(环)．

［例 11］　我们知道，对于以实数为元素的矩阵的和及积，

$\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{B}+\boldsymbol{A}$, $^\forall\boldsymbol{A}$, $\boldsymbol{B}\in R^{m\times n}$(加法的交换律)

$(\boldsymbol{A}+\boldsymbol{B})+\boldsymbol{C}=\boldsymbol{A}+(\boldsymbol{B}+\boldsymbol{C})$, $^\forall\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}\in R^{m\times n}$(加法的结合律)

* 原文误为 $s=R$. ——译者注

$$(\boldsymbol{AB})\boldsymbol{C}=\boldsymbol{A}(\boldsymbol{BC}),\ {}^\forall\boldsymbol{A}\in R^{m\times p},\ {}^\forall\boldsymbol{B}\in{}^{p\times q},\ {}^\forall\boldsymbol{C}\in{}^{q\times n}(\text{乘法的结合律})$$

$$\boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C})=\boldsymbol{AB}+\boldsymbol{AC},\ {}^\forall\boldsymbol{A}\in R^{m\times p},\ {}^\forall\boldsymbol{B},\ \boldsymbol{C}\in R^{p\times n}(\text{分配律}(1))$$

$$(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{C}=\boldsymbol{AC}+\boldsymbol{BC},\ {}^\forall\boldsymbol{A},\ \boldsymbol{B}\in R^{m\times p},\ {}^\forall\boldsymbol{C}\in R^{p\times n}(\text{分配律}(2))$$

成立．只要对矩阵的元素有限次地使用加法的交换律，加法、乘法的结合律及分配律就可以证明这些性质．因此，这些性质不只限于以实数为元素的矩阵，对于以任意环（例如整数环、有理数体、多项式环、有理函数体、实轴上连续函数环、有限体）的元为元素的矩阵均成立．

［例 12］ 大家知道，以 $x_k(k=1,\ 2,\ \cdots,\ n)$ 为未知数，具有实系数 a_{ij}、b_i 的线性代数方程式

$$\sum_{j=1}^{n}a_{ij}x_j=b_i\quad i=1,\ 2,\ \cdots,\ n$$

当系数矩阵 $\boldsymbol{A}=[a_{ij}]=n\times n$ 是正则矩阵（$\boldsymbol{A}$ 的逆矩阵存在）时，具有唯一解，可以用所谓的克兰姆公式给出实数解．实际上，克兰姆公式仅仅是有限次利用四则运算组合成的．当 a_{ij}，b_i 是任意体（例如有理函数体）的元时也可用这个公式，给出属于该体的解．

［例 13］ 在矩阵理论中有所谓的凯莱-哈密顿（Cayley-Hamilton）定理，其内容是："任意复数元素方阵满足其特征方程式"．因该定理的证明过程由有限次四则运算组成，故对于任意体的方阵均成立．

但是另一方面，当然也要注意由于体的不同而产生的区别．例如，一般不能保证矩阵的特征值作为和矩阵的元素所在的同一个体的元求出．考虑到实数矩阵的特征值常不限于实数，这一点就清楚了．

基于上述理由，因此今后讲述矩阵理论中的各种结果时必须注意对于其它体是否都成立，或必须考虑和其它体的区别．

B-0-2　向量空间和矩阵

让我们再从和前一章不同的角度，对矩阵再稍微深入地讨论一下.

实数 ξ_1 可以在一维的直线上，用连结坐标原点和坐标为 ξ_1 的点得到的一维向量表示，而复数 $\xi_1+j\xi_2$ 可以在所谓的复平面上，用连结坐标原点和坐标为 (ξ_1, ξ_2) 的 2 维向量表示. 大家在经典控制理论中已经掌握了向量轨迹等概念，故对此特别熟悉.

一般，若平面上设有坐标系统，则连结坐标原点和坐标为 (ξ_1, ξ_2) 的点的向量可以用 (ξ_1, ξ_2) 表示. 同样，(ξ_1, ξ_2, ξ_3) 可以表示 3 维空间中的向量. 尽管如此，那么更为一般，$(\xi_1, \xi_2, \cdots, \xi_n)$ 进而 $m\times n$ 矩阵 $[\xi_{ij}]$ 等等又如何呢？我们能直接认识的是直线、平面、到立体空间为止，但是，作为将它们推广的抽象概念，可以定义向量空间. 在经典控制理论及力学中讨论向量时，向量的和定义为，以各向量对应分量的和为分量的向量；实数 α 和向量相乘定义为，该向量的各分量都增大 α 倍. 那么，关于加法和纯量倍的几个规律成立. 将这些规律推广可作出如下定义.

［定义 1］　当集合 V 满足下列公理时，称 V 为向量空间 (Vector space)，V 的元称为向量.

a）对于 $\forall \boldsymbol{x}, \boldsymbol{y}\in V$, $\boldsymbol{x}+\boldsymbol{y}$ 有意义，V 是加群.

b）当体 F 给定时，对于 $\forall\alpha\in F$, $\forall\boldsymbol{x}\in V$，纯量倍 $\alpha\boldsymbol{x}$ 有意义，而且 $\alpha\boldsymbol{x}\in V$ 是唯一确定的.

c）$\alpha(\beta\boldsymbol{x})=(\alpha\beta)\boldsymbol{x}$, $\forall\boldsymbol{x}\in V$, $\forall\alpha, \beta\in F$.

d）对于 F 的单位元 1, $1\boldsymbol{x}=\boldsymbol{x}$, $\forall\boldsymbol{x}\in V$.

e）$\alpha(\boldsymbol{x}+\boldsymbol{y})=\alpha\boldsymbol{x}+\alpha\boldsymbol{y}$, $\forall\boldsymbol{x}^{*}, \boldsymbol{y}\in V$, $\forall\alpha\in F$.

f）$(\alpha+\beta)\boldsymbol{x}=\alpha\boldsymbol{x}+\beta\boldsymbol{x}$, $\forall\boldsymbol{x}\in V$, $\forall\alpha, \beta\in F$.

对于数 α 所属的体 F，例 5 中的实数体 R、复数体 C 等用得最多. 各体常分别称为实空间、复空间，等等. 此外，也有写成 $V(F)$，表示体 F 上的向量空间.

（问题 1）　因为向量空间 $V(F)$ 是加群，所以有关于加法的单位元（零元）. 设其为 $\boldsymbol{\theta}$，称为零向量. 此外，设 V 的元 x 关于加法的逆元为 $-\boldsymbol{x}$，即 $\boldsymbol{x}+(-\boldsymbol{x})=\boldsymbol{\theta}$，试证明

a）$\boldsymbol{\theta}$是唯一确定的

b）$0\cdot\boldsymbol{x}=\boldsymbol{\theta}$, $\forall\boldsymbol{x}\in V(F)$

c）$\alpha\boldsymbol{\theta}=\boldsymbol{\theta}$, $\forall\alpha\in F$

d）$(-1)\cdot\boldsymbol{x}=-\boldsymbol{x}$, $\forall\boldsymbol{x}\in V(F)$

式中 0 和 1 分别为各体关于加法和乘法的单位元.

［例 1］　平面上向量 (ξ_1, ξ_2) 的全体，是关于上述加法和纯量倍的实向量空间.

［例 2］　一般，n 个实数（复数）组 $\boldsymbol{x}=(\xi_1, \cdots, \xi_n)$ 的全体，若可以用 P-0-1 中 (3) 式的加法和 (5) 式的纯量倍，则为实向量空间，用 $R^n(C^n)$ 表示.

---

* 原文误印成 $\forall\boldsymbol{y}, \boldsymbol{y}\in V$. ——译者注

例 2 是平面向量的直接引深．但是，向量空间的概念不限于此，它具有更广泛的概括意义．例如

［例 3］ 体 F 自身是 F 上的向量空间．

［例 4］ $P(s, F)$是 F 上的向量空间．

［例 5］ 实轴上定义的连续实函数的全体，是在 R 上的关于普通加法和纯量倍的向量空间．

［例 6］ 设 V、W 是同一个体 F 上的向量空间，将 $\boldsymbol{x}\in V$，$\boldsymbol{y}\in W$ 按顺序标以脚注，对于$(\boldsymbol{x}, \boldsymbol{y})$对，若

$$(\boldsymbol{x}_1, \boldsymbol{y}_1)+(\boldsymbol{x}_2, \boldsymbol{y}_2)\triangleq(\boldsymbol{x}_1+\boldsymbol{x}_2, \boldsymbol{y}_1+\boldsymbol{y}_2)$$
$$\alpha(\boldsymbol{x}, \boldsymbol{y})\triangleq(\alpha\boldsymbol{x}, \alpha\boldsymbol{y})$$

有意义，则$(\boldsymbol{x}, \boldsymbol{y})$的全体形成 F 上的向量空间，叫做 V 和 W 的直积，用 $V\times W$ 表示．对于三个以上的向量空间，将直积推广到一般情况是很容易的．特别是，常把 $V\times V\times\cdots\times V$ (n 次)写成 V^n．

根据例 3 和例 6，在任意体 F 中每次取 n 个元按着次序排列，所得排列的全体是 F 上的向量空间，写成 F^n．例 2 中的 R^n 和 C^n 是其中的一个例子．

［例 7］ 设 F 上 $m\times n$ 矩阵的全体为 $F^{m\times n}$，则 $F^{m\times n}$ 是体 F 上关于 P-0-1(3)，(5) 式运算的向量空间．

也就是说，可以看作向量是矩阵的一个特点．

下面再稍微讲一下以后经常要用到的向量的线性独立及线性相关的概念．

［定义 2］ 当 $\boldsymbol{x}_i\in V(F)$ $(i=1, \cdots, n)$全不为 0 时，若有某些 $\alpha_i\in F$ $(i=1, \cdots, n)$ 满足

$$\alpha_1\boldsymbol{x}_1+\cdots+\alpha_n\boldsymbol{x}_n=\boldsymbol{\theta} \tag{1}$$

则 $\boldsymbol{x}_1$ $\cdots$, $\boldsymbol{x}_n$ 称为线性相关．若(1)式仅在 $\alpha_1=\alpha_2=\cdots=\alpha_n=0$ 时才成立，则 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$ 称为线性独立．

［例 8］ 现来看一下 R^3(例 2)，对于

$$\boldsymbol{x}_1\triangleq\begin{bmatrix}1\\0\\0\end{bmatrix},\quad \boldsymbol{x}_2\triangleq\begin{bmatrix}2\\1\\0\end{bmatrix},\quad \boldsymbol{x}_3\triangleq\begin{bmatrix}3\\1\\2\end{bmatrix}$$

立即可以看到，仅当 $\alpha_1=\alpha_2=\alpha_3=0$ 时(1)式才成立，所以 $\boldsymbol{x}_1$、$\boldsymbol{x}_2$、$\boldsymbol{x}_3$ 是线性独立的．若再加上

$$\boldsymbol{x}_4\triangleq\begin{bmatrix}1\\1\\1\end{bmatrix}$$

因为当 $\alpha_1=-\dfrac{3}{2}$，$\alpha_2=\dfrac{1}{2}$，$\alpha_3=\dfrac{1}{2}$，$\alpha_4=-1$ 时(1)式成立，则 $\boldsymbol{x}_1$、$\boldsymbol{x}_2$、$\boldsymbol{x}_3$、$\boldsymbol{x}_4$ 线性相关．

［例 9］ 因为 $Q(s, R)$ 是体（见前面例 7)，根据例 3，则 $Q(s, R)$ 是其自身上的向量空间．此外，根据例 6，$Q(s, R)\times Q(s, R)$ 也是 $Q(s, R)$ 上的向量空间．对此，试看下列两个向量

$$\boldsymbol{x}_1 \triangleq \begin{bmatrix} \dfrac{1}{s+1} \\ \dfrac{1}{s+2} \end{bmatrix}, \quad \boldsymbol{x}_2 \triangleq \begin{bmatrix} \dfrac{s+2}{(s+1)(s+5)} \\ \dfrac{1}{s+5} \end{bmatrix}$$

对于 $\alpha_1, \alpha_2 \in Q(s, R)$，因选取

$$\alpha_1 = -1, \quad \alpha_2 = \frac{s+5}{s+2}$$

时(1)式成立,则 $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ 线性相关.

［例 10］ 容易证明, $Q(s, R)$ 形成体 R 上的向量空间,则 $Q(s, R) \times Q(s, R)$ 也是 R 上的向量空间. 试看一下和例 9* 相同的 $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, 对于 $\alpha_1, \alpha_2 \in R$, 仅当 $\alpha_1 = \alpha_2 = 0$ 时 (1) 式才成立,因此 $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ 线性独立.

如例 9、例 10 所示,即使是同一组向量,但若定义向量空间时的体不同,则其可能线性相关,也可能线性独立.

［例 11］ $1, t, t^2, \cdots, t^n$ 都是例 5 中向量空间的元,而且它们是线性独立的 (例 5 中向量空间中的零向量 θ, 是"对于所有的 t 恒为 0 的函数"的向量).

［定义 3］ 向量空间的线性独立向量的最大个数,叫做向量空间的维数.

［例 12］ R^n 的维数是 n (以后证明).

［例 13］ 在例 5 的向量空间中, $1, t, t^2, \cdots$, 是线性独立的,则该向量空间为无限维向量空间.

现在,再回到我们所能直接理解的平面和立体空间. 对于该空间的一切向量,试看对应于某些向量的操作. 例如

1) 将平面向量逆时针转过 90°, 将其长度延长到原来的 2 倍;

2) 对于在 ξ_1、ξ_2、ξ_3 空间中的向量 $\boldsymbol{x}$, 从该向量的矢端作垂直于 $\xi_3=0$ 平面的垂线,连结该垂线的垂足和坐标原点,便得到对应的向量 $\boldsymbol{y}$ (从与 $\xi_3=0$ 平面垂直方向射来平行光线时, 向量 $\boldsymbol{x}$ 在 $\xi_3=0$ 平面上的投影是 $\boldsymbol{y}$. 因此, $\boldsymbol{y}$ 叫做 $\boldsymbol{x}$ 在 $\xi_3=0$ 平面上的正交投影,等等).

一般将这种对应关系称为映射,关于在向量空间中的映射,引入下列定义.

［定义 4］ 由某向量空间 $V(F)$ 到具有相同体的另一向量空间 $V'(F)$ 的映射 T, 当满足下列公理时,称为线性变换(linear transformation).

a) $T(\boldsymbol{x}+\boldsymbol{y}) = T\boldsymbol{x}+T\boldsymbol{y}, \ ^\forall\boldsymbol{x}, \boldsymbol{y} \in V(F)$

b) $T(\alpha\boldsymbol{x}) = \alpha T\boldsymbol{x}, \ ^\forall\boldsymbol{x} \in V(F), \ ^\forall\alpha \in F$

［例 14］ R^n 的向量乘以 $R^{m \times n}$ 的矩阵, 便得到 R^m 的向量, 它是由 R^n 到 R^m 的线性变换.

在例 14 中谈到, 乘以矩阵是线性变换. 现在再稍微谈一下相反的问题, 即一般任意的线性变换能否用矩阵表示.

首先,容易看出前面在 1)、2) 中所讲的映射都是线性变换. 而且, 实际上 1)、2) 的映射可表示如下

1) $\boldsymbol{y} = A\boldsymbol{x}, \quad A = \begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}$

* 原文误印成例 10. ——译者注

2) $\boldsymbol{y}=\boldsymbol{A}\boldsymbol{x}$, $\quad \boldsymbol{A}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$,

可以很简单地用在 $\boldsymbol{x}$ 上乘以某矩阵 $\boldsymbol{A}$ 得出. 也就是说, 1)、2) 的线性变换能够用矩阵表示, 其实该结论在更一般的情况下也成立. 因为要再讲一些有关内容后才能对它进行更详细的讨论,所以这里先在上述例子的基础上大概谈谈以下的内容. "在有限维向量空间内, 若规定坐标系统, 则线性变换能用矩阵表示". 这就说明矩阵具有表示线性变换的特性,这是矩阵的第二个特点.

原讲座写作计划

原讲座计划写 I, II, III, IV 四个部分(本书译出的是其中的第 I 部分).从各个不同角度阐述矩阵理论.

在第 I 部分中,将矩阵看作体 F (主要是 R 或 C) 的元的排列,阐明矩阵的运算及分析方面的内容.因此,主要内容是,矩阵运算的注意事项,基本变换及其应用,与矩阵特征值有关的事项,矩阵分析,矩阵函数,矩阵方程等等.

在第 II 部分中,将矩阵考虑成从有限维向量空间 $V(F)$ 到 $W(F)$ 的线性变换表示.因此,要讨论向量空间和其部分空间的基底及线性变换的值域、零空间、不变部分空间等有关的内容.利用这些概念,将第 I 部分所讨论的内容再一次从几何学的观点加以讨论,这是第 II 部分的目的.

在第 III 部分中,以有理函数体上的矩阵为对象.说明基本变换及其应用,矩阵方程等.

在第 IV 部分中,讨论有限体上的矩阵,这是分析线性顺序系统所必要的基本理论.

第 I~IV 部分内容,是为了叙述方便而划分的,而研究某些自动控制中的问题时,往往需要其中两个部分或更多部分的内容.例如,试考虑状态方程式用

$$\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{u} \tag{1}$$

表示的常系数线性系统的可控性.这时若设 $\boldsymbol{A}\in R^{n\times n}$, $\boldsymbol{B}\in R^{n\times r}$ (即 $\boldsymbol{A}$, $\boldsymbol{B}$ 分别为实数体上的 $n\times n$ 及$n\times r$矩阵),众所周知,可控性的充分与必要条件是

$$\operatorname{rank}[\boldsymbol{B},\ \boldsymbol{A}\boldsymbol{B},\ \cdots,\ \boldsymbol{A}^{n-1}\boldsymbol{B}]=n \tag{2}$$

这是利用第 I 部分内容给出的条件.第 II 部分可以给出与其完全等价的条件,例如可以叙述为:"包含 $\boldsymbol{B}$ 的值域的 $\boldsymbol{A}$ 的最小不变部分空间和系统的状态空间一致".

而且,考虑到状态变量 $\boldsymbol{x}$ 与输入量 $\boldsymbol{u}$ 之间的传递矩阵

$$\boldsymbol{G}_x(s)=[s\boldsymbol{I}-\boldsymbol{A}]^{-1}\boldsymbol{B},$$

则可控性的充分和必要条件(如本书后面部分所述)变为:"$\boldsymbol{G}_x(s)$ 的特征方程式的次数是n",它与第 III 部分内容有关.

其次,试考虑状态方程用下列差分方程

$$\boldsymbol{x}(k+1)=\boldsymbol{A}\boldsymbol{x}(k)+\boldsymbol{B}\boldsymbol{u}(k),\quad k=0,\ 1,\ 2,\ \cdots \tag{3}$$

表示的常系数线性顺序系统.式中状态变量 $\boldsymbol{x}$ 的分量和矩阵 $\boldsymbol{A}(=n\times n)$, $\boldsymbol{B}(=n\times r)$的元素都是有限体 $GF(p)$ 的元,认为它服从该体上的运算(即 modp).这时,系统可控性的充分和必要条件可用和(2)式相同的形式给出,其推导过程直接用到第 IV 部分的内容.

这样,在讨论系统的可控性时,就必然涉及到第 I~IV 部分的内容.实际上,对于表 1 所列的控制理论中的许多基本问题也是一样.通过以后各部分中关于控制问题的讨论,这一点将会逐渐清楚.

补注

在加群 G 中,我们将 $\alpha\in G$ 的逆元表示成 $-\alpha$, 把 $\beta\in G$ 和 $-\alpha$ 的和记为 $\beta-\alpha$,

$$\beta-\alpha \triangleq \beta+(-\alpha)$$

但是“-1”这个记号具有两种意义.

a) 是元 1 的加法逆元.

b) 是对于整数(实数或复数)的“负 1”.

当然,在环 Z, 体 R, C 等中, a)中(-1)的意义和 b)中一致,没有区别的必要. 但是在 $GF(2)$中, a)中的(-1)是指元 1, 而对于 $GF(2)$中的元来说, b) 中的(-1)是不存在的. 同样,在 $GF(3)$中,根据 a) 的意义

$$-1=2$$
$$-2=3$$

成立. 而根据 b)的意义, -1、-2 不是 $GF(3)$的元.

设矩阵 $\boldsymbol{A}\in\tilde{R}^{m\times n}$ 的元素为 $a_{ij}\in\tilde{R}$, 以 a_{ij} 在 $\tilde{R}$ 中的加法逆元$(-a_{ij})$为 (i, j) 元素的矩阵,称为 $-\boldsymbol{A}$.

关于矩阵和 $\tilde{R}^{m\times n}$ 构成加群,则易见 $\boldsymbol{A}$ 的加法逆元是上面定义的$-\boldsymbol{A}$.

[例]

对于体 R 上的矩阵

$$\boldsymbol{A}=\begin{bmatrix}1 & 1 & 0\\ 0 & 0 & 1\end{bmatrix},\quad -\boldsymbol{A}=\begin{bmatrix}-1 & -1 & 0\\ 0 & 0 & -1\end{bmatrix}$$

对于体 $GF(2)$上的矩阵

$$\boldsymbol{A}=\begin{bmatrix}1 & 1 & 0\\ 0 & 0 & 1\end{bmatrix},\quad -\boldsymbol{A}=\begin{bmatrix}1 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$

对于体 $GF(3)$上的矩阵

$$\boldsymbol{A}=\begin{bmatrix}1 & 1 & 0\\ 0 & 0 & 1\end{bmatrix},\quad -\boldsymbol{A}=\begin{bmatrix}2 & 2 & 0\\ 0 & 0 & 2\end{bmatrix}$$

P-I-1 矩阵的种类

上面已经指出，自动控制中所遇到矩阵的元素不只限于实数，而涉及到复数、多项式、有理函数、变量 t 的连续函数等等. 为了系统地进行讨论，应用环、体的概念是很方便的，以后还要借助于这些概念尽可能进行更普遍的讨论. 但是大家对这些概念比较生疏. 在感到抽象的情况下可以这样来理解，谈到"体 F 上的矩阵"时，可以认为是在讨论以实数为元素的矩阵、以复数为元素的矩阵、以有理函数为元素的矩阵的共同性质. 在讲到"环 $\widetilde{R}$ 上的矩阵"时，除上述矩阵外，例如可以认为对于以整数为元素以及以多项式为元素的矩阵亦均成立.

一般的环 $\widetilde{R}$ 不一定存在乘法单位元，而且乘法的交换律也不一定成立. 但是，现在所讨论的对象以多项式环为代表，因为它是具有乘法单位元的可换环，所以本章内我们假定 $\widetilde{R}$ 是具有乘法单位元的可换环.

$\widetilde{R}$ 上 n 阶方阵的全体 $\widetilde{R}^{n\times n}$ 构成环（参看 B-0-1，例 3)，则由乘法的结合律，对于 $\boldsymbol{A}\in\widetilde{R}^{n\times n}$，乘积

$$\underbrace{\boldsymbol{A}\boldsymbol{A}\cdots\boldsymbol{A}}_{k\text{个}}$$

与运算次序无关，是唯一确定的，可以写成 $\boldsymbol{A}^k$，称为 $\boldsymbol{A}$ 的 k 次方幂.

$\boldsymbol{A}^2=\boldsymbol{A}$ 的矩阵称为幂等矩阵. 对于某个自然数 k_0，$\boldsymbol{A}^{k_0}=\boldsymbol{0}_{m\times n}$ 的矩阵称为幂零矩阵.

$\widetilde{R}$ 上 $m\times n$ 矩阵的全体用 $\widetilde{R}^{m\times n}$ 表示，设 $\boldsymbol{A}\in\widetilde{R}^{m\times n}$，以 $\boldsymbol{A}$ 的 (j, i) 元作为 (i, j) 元素的矩阵 * $[a_{ji}]\in\widetilde{R}^{m\times n}$，称为 $\boldsymbol{A}$ 的转置矩阵，用 $\boldsymbol{A}^T$ 表示. 关于转置矩阵，下列公式成立

(i) $(\boldsymbol{A}^T)^T=\boldsymbol{A}$, ${}^\forall\boldsymbol{A}\in\widetilde{R}^{m\times n}$

(ii) $(\boldsymbol{A}+\boldsymbol{B})^T=\boldsymbol{A}^T+\boldsymbol{B}^T$, ${}^\forall\boldsymbol{A}, \boldsymbol{B}\in\widetilde{R}^{m\times n}$

(iii) $(\boldsymbol{A}\boldsymbol{B})^T=\boldsymbol{B}^T\boldsymbol{A}^T$, ${}^\forall\boldsymbol{A}\in\widetilde{R}^{m\times p}$, ${}^\forall\boldsymbol{B}\in\widetilde{R}^{p\times n}$

(iv) $(\alpha\boldsymbol{A})^T=\alpha\boldsymbol{A}^T$, ${}^\forall\boldsymbol{A}\in\widetilde{R}^{m\times n}$, ${}^\forall\alpha\in\widetilde{R}$

对于方阵 $\boldsymbol{A}\in\widetilde{R}^{n\times n}$，当 $\boldsymbol{A}^T=\boldsymbol{A}$ 时，$\boldsymbol{A}$ 叫做对称阵，当 $\boldsymbol{A}^T=-\boldsymbol{A}$ 时，$\boldsymbol{A}$ 称为斜对称阵. 关于对称性有如下性质

(i) 对于任意 $\boldsymbol{A}\in\widetilde{R}^{m\times n}$，$\boldsymbol{A}\boldsymbol{A}^T$、$\boldsymbol{A}^T\boldsymbol{A}$ 总是可以定义的，并且都是对称阵.

(ii) 而且仅当 $\boldsymbol{A}$ 是方阵时，$\boldsymbol{A}+\boldsymbol{A}^T$、$\boldsymbol{A}-\boldsymbol{A}^T$ 才有意义，$\boldsymbol{A}+\boldsymbol{A}^T$ 为对称阵，$\boldsymbol{A}-\boldsymbol{A}^T$ 为斜对称阵.

(iii) 若 $\boldsymbol{A}\in\widetilde{R}^{n\times n}$ 为对称阵，则

a) $\alpha\boldsymbol{A}$ 也是对称阵，${}^\forall\alpha\in\widetilde{R}$.

b) $\boldsymbol{A}\boldsymbol{A}^T=\boldsymbol{A}^T\boldsymbol{A}$.

c) $\boldsymbol{A}^k(k=1, 2, \cdots)$ 也是对称阵.

(iv) 另一方面，若 $\boldsymbol{A}\in\widetilde{R}^{n\times n}$ 为斜对称阵，则

a) $\alpha\boldsymbol{A}$ 也是斜对称阵，${}^\forall\alpha\in\widetilde{R}$

* 即将矩阵的行和列互换所得到的矩阵. ——译者注

b）$AA^T=A^TA$

c）当 k 为偶数时，$A^k(k=1, 2, \cdots)$ 为对称阵，当 k 为奇数时，A^k 为斜对称阵.

(v) 若 $A, B\in \tilde{R}^{n\times n}$ 均为对称阵，则

a）$A+B$ 也是对称阵

b）若 $AB=BA$，则 AB 也是对称阵.

下面我们限定 A 是实数体 R 上的矩阵，那末

(vi) 若 $A\in R^{n\times n}$ 为斜对称阵，则其主对线元素均为 0.

(vii) $A\in R^{n\times n}$ 可以分解成对称阵和斜对称阵之和，即

$$A=\underbrace{\frac{1}{2}(A+A^T)}_{\text{对称}}+\underbrace{\frac{1}{2}(A-A^T)}_{\text{斜对称}} \tag{1}$$

这两个性质不限于 R，对于 C, $P(s, R)$, $Q(s, R)$ 等上的方阵也同样成立. 但是，并不是在所有情况下都是正确的，例如，对于 $GF(2)$ 上的任意方阵，无论如何也不会成立.

［问题 1］ 试证明(vi)，(vii)在 $GF(2)$ 上不成立.

其次，把 A 作为复数体 C 上的矩阵(不限于 C，在 $P(s, C)$，$Q(s, C)$ 的情况下也是一样). 将 $A\in C^{m\times n}$ 的各元素 a_{ij} 用其共轭复数 $\bar{a}_{ij}$ 置换，得到的矩阵 $[\bar{a}_{ij}]\in C^{m\times n}$ 写为 $\bar{A}$，称为 A 的共轭阵. 并且把$(\bar{A})^T\in C^{m\times n}$ 写成 A^*. 对于 A, A^* 有下列公式:

(i) $A^*\triangleq(\bar{A})^T=(\overline{A^T})$ $\quad {}^\forall A\in C^{m\times n}$

(ii) $\overline{(A+B)}=\bar{A}+\bar{B}$ $\quad {}^\forall A, B\in C^{m\times n}$

(iii) $\overline{AB}=\bar{A}\cdot\bar{B}$ $\quad {}^\forall A\in C^{m\times p}$, ${}^\forall B\in C^{p\times n}$

(iv) $\overline{(\alpha A)}=\bar{\alpha}\bar{A}$ $\quad {}^\forall A\in C^{m\times n}$, ${}^\forall \alpha\in C$

［问题 2］ 试求$(A+B)^*$，$(AB)^*$，$(\alpha A)^*$.

对于方阵，当 $A^*=A$ 时，A 称为埃尔米特阵，$A^*=-A$ 时，称为斜埃尔米特阵.

前边所述的对称性与埃尔米特性很相似，从而很容易得出对应的埃尔米特性质.

［问题 3］ 对照前边例举的对称性诸性质，试叙述对应的埃尔米特性质.

除此之外，作为埃尔米特性固有的性质，任意的 $A\in C^{n\times n}$ 可以表示成

$$A=A_1+jA_2 \quad A_1, A_2\in C^{n\times n} \tag{2}$$

式中 A_1, A_2 都是埃尔米特矩阵.

在 $A\in \tilde{R}^{n\times n}$ 中，$a_{ij}(j>i)$ 全为 0 的矩阵称为下三角形矩阵. 相反，$a_{ij}(j<i)$ 全为 0 的矩阵称为上三角形矩阵. $a_{ij}(j\neq i)$ 全为 0 的矩阵称为对角线矩阵. 在对角线矩阵中，对角线上元素全为 1 的矩阵称为单位矩阵 I_n.

在方阵 A 中，1 在各行各列中仅出现一次，而矩阵中其它元素全为 0 的矩阵叫做顺列矩阵. 在 $n\times 1$ 阶矩阵中，仅第 i 行元素为 1，其它元素全为 0 的矩阵写成 e_i，即

$$e_i=\begin{bmatrix}0\\ \vdots\\ 0\\ 1\\ 0\\ \vdots\\ 0\end{bmatrix}\begin{matrix}\\ \\ \\ \leftarrow\text{第 } i \text{ 行元素}\\ \\ \\ \\ \end{matrix} \tag{3}$$

使用该记号则 n 阶单位阵可写成

$$I_n=[e_1e_2\cdots e_n]=\begin{bmatrix} e_1^T \\ e_2^T \\ \cdots \\ e_n^T \end{bmatrix} \tag{4}$$

而且任意 $n\times n$ 阶顺列矩阵 P_n 可以表示成

$$P_n=[e_{i_1}e_{i_2}\cdots e_{i_n}]=\begin{bmatrix} e_{j_1}^T \\ e_{j_2}^T \\ \vdots \\ e_{j_n}^T \end{bmatrix} \tag{5}$$

式中 $(i_1, i_2, \cdots, i_n)$，$(j_1, j_2, \cdots, j_n)$ 分别为将整数数组 $(1, 2, \cdots, n)$ 排列得到的排列之一. 任意 $m\times n$ 矩阵 A 左乘以 $m\times m$ 阶顺列矩阵 P_m，就相当于分别将 A 的第 j_1 行移到第 1 行，第 j_2 行移到第 2 行，……，第 j_m 行移到第 m 行，即可用"行的替换"表示. A 右乘以 $n\times n$ 阶顺列矩阵 P_n 相当于分别将 A 的第 i_1 列移到第 1 列，……，第 i_n 列移到第 n 列，即可用"列的替换"表示.

下三角形矩阵、上三角形矩阵、对角形矩阵等各自的和以及纯量倍，仍然分别为下三角形矩阵、上三角形矩阵、对角形矩阵等. 下三角形矩阵、上三角形矩阵、对角形矩阵、单位矩阵以及顺列矩阵等各自的积，仍然分别为下三角形矩阵、上三角形矩阵、对角形矩阵、单位矩阵以及顺列矩阵等.

在矩阵 A 中画若干条水平线和垂直线，将它分成 n 块，例如

$$\begin{bmatrix} 1 & 2 & -1 & 0 \\ 1 & 8 & 7 & 6 \\ 0 & 1 & 3 & 5 \\ 1 & 0 & 0 & 0 \end{bmatrix}=\left[\begin{array}{ccc|c} 1 & 2 & -1 & 0 \\ 1 & 8 & 7 & 6 \\ \hline 0 & 1 & 3 & 5 \\ 1 & 0 & 0 & 0 \end{array}\right]=\left[\begin{array}{cc|cc} 1 & 2 & -1 & 0 \\ 1 & 8 & 7 & 6 \\ \hline 0 & 1 & 3 & 5 \\ 1 & 0 & 0 & 0 \end{array}\right]$$

划成的各块是比 A 小的矩阵，A 变成以其小矩阵为元素的矩阵. 这种小矩阵称为矩阵 A 的子块. 由上例可见，同一矩阵可划分成不同的子块. 但是无论如何划分，显然它们必须满足下列两个条件

(i) 各子块的行数、列数都是非负整数

(ii) 所有子块行数之和、列数之和，分别等于矩阵的行数、列数.

分块矩阵可表示如下

$$A=\begin{bmatrix} A_{11} & \cdots & A_{1p} \\ \vdots & & \vdots \\ A_{q1} & \cdots & A_{qp} \end{bmatrix} \tag{6}$$

当 $q=p$ 时，A_{ii} 均为方阵. 而且，当 $A_{ij}(j>i)$ 均为零阵时，A 称为分块下三角形矩阵. 同样，也可以定义分块上三角形矩阵、分块对角形矩阵. 特别是，分块对角形矩阵

$$A=\begin{bmatrix} A_{11} & & & \mathbf{0} \\ & A_{22} & & \\ & & \ddots & \\ \mathbf{0} & & & A_{pp} \end{bmatrix} \tag{7}$$

也称为 $A_{11}, A_{22}, \cdots, A_{pp}$ 的直和，常写成下列形式

$$A=A_{11}\dotplus A_{22}\dotplus\cdots\dotplus A_{pp} \tag{8}$$

方阵 $\boldsymbol{A}$ 为对称(斜对称)阵的充分和必要条件是，当 $\boldsymbol{A}_{ii}$ 都划分成方阵时，$\boldsymbol{A}_{ii}$ 是对称（斜对称)阵,对于 $j\neq i$, $\boldsymbol{A}_{ij}=\boldsymbol{A}_{ji}^T(\boldsymbol{A}_{ij}=-\boldsymbol{A}_{ji}^T)$ 成立.

当 $\boldsymbol{A}\in R^{n\times n}$ 为实对称阵，$\boldsymbol{x}\in R^{n\times 1}$ 时,乘积 $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$ 叫做二次型,二次型是实数.

[问题 4] 当 $\boldsymbol{A}\in R^{n\times n}$ 不一定是对称阵,而是一般矩阵时,试证明

$$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}=\frac{1}{2}\boldsymbol{x}^T(\boldsymbol{A}+\boldsymbol{A}^T)\boldsymbol{x} \tag{9}$$

设 $\boldsymbol{A}\in R^{m\times n}$, $\boldsymbol{y}\in R^{m\times 1}$, $\boldsymbol{x}\in R^{n\times 1}$，则乘积 $\boldsymbol{y}^T\boldsymbol{A}\boldsymbol{x}$ 称为双一次型. 双一次型也是实数.

设 $\boldsymbol{A}\in C^{n\times n}$ 为埃尔米特阵，$\boldsymbol{x}\in C^{n\times 1}$，则乘积 $\boldsymbol{x}^*\boldsymbol{A}\boldsymbol{x}$ 称为埃尔米特型.

[问题 5] 试证明埃尔米特型是实数.

根据二次型的符号,实对称阵可分类如下:

若 $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}>0$, $\forall \boldsymbol{x}\neq \boldsymbol{0}_{n\times 1}$，则 $\boldsymbol{A}$ 称为正定阵

$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}\geqslant 0$, $\forall \boldsymbol{x}$	$\boldsymbol{A}$ 称为准正定阵 （或称为正半定阵)
$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}\leqslant 0$, $\forall \boldsymbol{x}$	$\boldsymbol{A}$ 称为准负定阵 （或称为负半定阵)
$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}<0$, $\forall \boldsymbol{x}\neq \boldsymbol{0}_{n\times 1}$	$\boldsymbol{A}$ 称为负定阵

对于埃尔米特型也是一样.

设 F 为任意体,对于 $\boldsymbol{A}\in F^{n\times n}$，若有满足 $\boldsymbol{AX}=\boldsymbol{I}_n$, $\boldsymbol{XA}=\boldsymbol{I}_n$ 的 $\boldsymbol{X}\in F^{n\times n}$ 存在，而且 $\boldsymbol{X}$ 唯一时,则 $\boldsymbol{X}$ 称为 $\boldsymbol{A}$ 的逆矩阵,写成 $\boldsymbol{A}^{-1}$.

对于逆矩阵

(i) 若 $\boldsymbol{A}^{-1}$ 存在,则 $(\boldsymbol{A}^{-1})^{-1}$ 必然存在且等于 $\boldsymbol{A}$.

(ii) 设 $\boldsymbol{A}$, $\boldsymbol{B}\in F^{n\times n}$，若 $\boldsymbol{A}^{-1}$, $\boldsymbol{B}^{-1}$ 存在,则 $(\boldsymbol{AB})^{-1}$ 亦存在,且 $(\boldsymbol{AB})^{-1}=\boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$.

(iii) 若 $\boldsymbol{A}^{-1}$ 存在,则 $(\boldsymbol{A}^T)^{-1}$, $(\overline{\boldsymbol{A}})^{-1}$ 亦存在,且

$$(\boldsymbol{A}^T)^{-1}=(\boldsymbol{A}^{-1})^T,\ (\boldsymbol{A})^{-1}=(\boldsymbol{A}^{-1})$$

具有逆矩阵的矩阵 $\boldsymbol{A}$ 叫做正则矩阵. 关于用什么方法判别一个矩阵是否是正则矩阵,以及正则矩阵的逆矩阵的求法,将在以后叙述.

设 $\boldsymbol{P}_1\in F^{m\times m}$, $\boldsymbol{P}_2\in F^{n\times n}$ 均为正则矩阵,则对于 $\boldsymbol{A}\in F^{m\times n}$，下式[1]

$$\boldsymbol{P}_1\boldsymbol{A}=\boldsymbol{0}_{m\times n}\Leftrightarrow \boldsymbol{A}=\boldsymbol{0}_{m\times n} \tag{10}$$

$$\boldsymbol{A}\boldsymbol{P}_2=\boldsymbol{0}_{m\times n}\Leftrightarrow \boldsymbol{A}=\boldsymbol{0}_{m\times n} \tag{11}$$

成立.

对于 $\boldsymbol{A}\in R^{n\times n}$，若存在 $\boldsymbol{A}^{-1}\in \widetilde{R}^{n\times n}$ 使得 $\boldsymbol{AA}^{-1}=\boldsymbol{A}^{-1}\boldsymbol{A}=\boldsymbol{I}_n$，则 $\boldsymbol{A}$ 称为幺模阵.

对于 $\boldsymbol{A}\in C^{n\times n}$，当 $\boldsymbol{A}^*\boldsymbol{A}=\boldsymbol{I}_n$ 时，换句话说，$\boldsymbol{A}^{-1}$ 存在和 $\boldsymbol{A}^*$ 一致时，则 $\boldsymbol{A}$ 称为酉阵. 当 $\boldsymbol{A}^T\boldsymbol{A}=\boldsymbol{I}_n$ 成立时，$\boldsymbol{A}$ 称为正交阵. 若 $\boldsymbol{A}=\boldsymbol{A}$，则酉性和正交性等价.

(i) 若 $\boldsymbol{A}$, $\boldsymbol{B}\in C^{n\times n}$ 均为酉(正交)阵,则 $\boldsymbol{AB}$ 也是酉(正交)阵.

(ii) 若 $\boldsymbol{A}\in C^{n\times n}$ 为酉阵,则 $\boldsymbol{A}^T$, $\overline{\boldsymbol{A}}$, $\boldsymbol{A}^*$, $\boldsymbol{A}^{-1}$ 均为酉阵. 若 $\boldsymbol{A}\in C^{n\times n}$ 是正交阵，则 $\boldsymbol{A}^T$, $\boldsymbol{A}^{-1}$ 也是正交阵.

[问题 6] 试证明顺序矩阵是正交阵.

[问题 7] 若对角形矩阵（分块对角形矩阵）的逆矩阵存在，试证明其逆矩阵也是对角形矩阵(分块对角形矩阵). 三角形矩阵的逆矩阵有什么性质?

1) $A\Rightarrow B$ 表示“如果 A 成立,则 B 亦成立”. 在 A 和 B 等价的情况下可用 $A\Leftrightarrow B$ 表示. 换句话说，$A\Rightarrow B$ 表示 A 是 B 的充分条件(即 B 是 A 的必要条件)，$A\Leftrightarrow B$ 表示 A 是 B 的充分必要条件.

P-I-2 行列式和迹

在本章内，我们来复习一下对于 $n\times n$ 方阵定义的两个纯量——行列式和迹.

行列式[9,10,19,20]

n 个整数(1, 2, …)换排的操作称为 n 文字置换*. n 文字置换共有 $n!$ 个. n 文字置换和反复进行若干次下列对换操作等价，每次只对换两个文字，而其它 n-2 个文字不动. 和偶数次对换等价的置换称为偶置换；和奇数次对换等价的置换称为奇置换.

例 $\{1,2,3\}\to\{2,1,3\}\to\{2,3,1\}\to\{3,2,1\}$

因而 $\{3,2,1\}$ 是奇置换

$\{2,3,1\}$ 是偶置换

对于体 F 上的 n 阶方阵 $\boldsymbol{A}=[a_{ij}]$,

$$\det\boldsymbol{A}\triangleq\sum_{p_j}^{n!}\pm a_{1j_1}a_{2j_2}\cdots a_{nj_n} \tag{1}$$

称为 $\boldsymbol{A}$ 的行列式 (determinant). 式中 Σ 表示 $n!$ 个置换的总和，符号 $\pm$ 按下列规则确定：将 $\{j_1, j_2, \cdots, j_n\}$ 看作由 $\{1, 2, \cdots, n\}$ 作出的排列，若

$\{1, 2, \cdots, n\}\to\{j_1, j_2, \cdots, j_n\}$ 是偶置换，则取+

$\{1, 2, \cdots, n\}\to\{j_1, j_2, \cdots, j_n\}$ 是奇置换，则取-

于是，若 $n=1$，则 $\det\boldsymbol{A}=a_{11}$，当 $n=2$ 时，

$$\det\boldsymbol{A}=a_{11}a_{22}-a_{12}a_{21} \tag{2}$$

n 阶 $(n\geqslant3)$ 方阵的行列式可以用 $(n-1)$ 阶方阵的行列式逐次进行计算.

设在 n 阵方阵中消去第 i 行和第 j 列 $(1\leqslant i, j\leqslant n)$ 后剩下的 $(n-1)$ 阶方阵为 $\boldsymbol{M}_{ij}$，这时

$$A_{ij}\triangleq\pm\det\boldsymbol{M}_{ij} \tag{3}$$

称为矩阵 $\boldsymbol{A}$ 中元素 a_{ij} 的代数余子式(cofactor). 至于 $\pm$ 符号，当 $i+j$ 为偶数时取 +，为奇数时取-. $\det\boldsymbol{A}$ 可用其代数余子式给出如下

$$\det\boldsymbol{A}=a_{1j}A_{1j}+a_{2j}A_{2j}+\cdots+a_{nj}A_{nj}\quad(1\leqslant j\leqslant n) \tag{4}$$

(4)式称为 $\boldsymbol{A}$ 的行列式对其第 j 列的展开式. 无论对那一列展开，都给出同样的行列式值.

例如，将 3×3 矩阵对第 3 列展开可得

$$\begin{aligned}\det\begin{bmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{bmatrix}&=a_{13}\det\begin{bmatrix}a_{21}&a_{22}\\a_{31}&a_{32}\end{bmatrix}-a_{23}\det\begin{bmatrix}a_{11}&a_{12}\\a_{31}&a_{32}\end{bmatrix}+a_{33}\det\begin{bmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{bmatrix}\\&=a_{13}(a_{21}a_{32}-a_{22}a_{31})-a_{23}(a_{11}a_{32}-a_{12}a_{31})+a_{33}(a_{11}a_{22}-a_{12}a_{21})\\&=a_{11}a_{22}a_{33}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}\end{aligned}$$

---

* 元素 $a_1, a_2, \cdots, a_n$ 的一个置换指的是对这 n 个元素施行这样一个变换，每一元素 a_i 被这 n 个元素中某一元素 a_j 所代替，并且不同的元素被不同的元素所代替. 在讨论 n 个元素的置换时，元素本身的性质对我们来说是不重要的，为了简单起见，这里用 n 个数码(1, 2, …)来代表这 n 个元素. ——译者注

将其对第 1 列或第 2 列展开时也可得到同样的结果.

行列式的基本性质[9,10,19,20]

由上述行列式的定义及展开式,可以得出下列基本性质:

(i) 矩阵转置后,其行列式值不变

$$\det\boldsymbol{A}=\det\boldsymbol{A}^T.$$

因此,关于行列式的行成立的性质,关于列也同样成立.

(ii) 若矩阵 $\boldsymbol{A}$ 一个行(或列)的元素全为 0, 则其行列式的值为 0.

(iii) 将行列式的任意一个列(或行)的各元素均乘以 α(α 是 F 的任意元), 则行列式的值也增大 α 倍. 对于列该性质可以表示如下. 首先将 $\boldsymbol{A}\in F^{n\times n}$ 写成

$$\boldsymbol{A}=[\boldsymbol{a}_1,\ \boldsymbol{a}_2,\ \cdots,\ \boldsymbol{a}_j,\ \cdots,\ \boldsymbol{a}_n]$$

若 $\boldsymbol{a}_j=\boldsymbol{A}$ 的第 j 列, 则

$$\det[\boldsymbol{a}_1,\ \cdots,\ \alpha\boldsymbol{a}_j,\ \cdots,\ \boldsymbol{a}_n]=\alpha\det\boldsymbol{A}$$

(iv) 当某一个列(例如第 j 列)为 $\boldsymbol{a}_j=\boldsymbol{b}_j+\boldsymbol{c}_j$ 时, 则 $\det\boldsymbol{A}$ 可表示成第 j 列为 $\boldsymbol{b}_j$ 的矩阵的行列式和第 j 列为 $\boldsymbol{c}_j$ 的矩阵的行列式之和, 即

$$\det[\boldsymbol{a}_1,\ \cdots,\ \underset{\text{第 }j\text{ 列}}{\underset{\uparrow}{\boldsymbol{b}_j+\boldsymbol{c}_j}},\ \cdots,\ \boldsymbol{a}_n]=\det[\boldsymbol{a}_1,\ \cdots,\ \underset{\text{第 }j\text{ 列}}{\underset{\uparrow}{\boldsymbol{b}_j}},\ \cdots,\ \boldsymbol{a}_n]+\det[\boldsymbol{a}_1,\ \cdots,\ \underset{\text{第 }j\text{ 列}}{\underset{\uparrow}{\boldsymbol{c}_j}},\ \cdots,\ \boldsymbol{a}_n]$$

成立. 关于行也存在同样的性质.

(v) 将某一列(行)乘以 α($\alpha\in F$) 后加到其它任意一列(行), 行列式的值不变. 即

$$\det[\boldsymbol{a}_1,\ \cdots,\ \boldsymbol{a}_n]=\det[\boldsymbol{a}_1,\ \cdots,\ \underset{\text{第 }i\text{ 列}}{\underset{\uparrow}{\boldsymbol{a}_i+\alpha\boldsymbol{a}_j}},\ \cdots,\ \underset{\text{第 }j\text{ 列}}{\underset{\uparrow}{\boldsymbol{a}_j}},\ \cdots,\ \boldsymbol{a}_n]$$

(vi) 交换任意两列(行), 行列式改变符号.

(vii) 若两列(行)相同, 则行列式的值为零.

(viii) 三角形矩阵的行列式值等于其主对角线元素之积, 即

$$\det\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}=\det\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}=a_{11}a_{22}\cdots a_{nn}$$

特别是, $\det\boldsymbol{I}=1$, $\det\boldsymbol{0}=0$.

矩阵的积和行列式

两个矩阵 $\boldsymbol{A}\in F^{n\times n}$ 和 $\boldsymbol{B}\in F^{n\times n}$ 的乘积 $\boldsymbol{AB}$ ($\boldsymbol{A}$ 和 $\boldsymbol{B}$ 不一定可交换) 或 $\boldsymbol{BA}$ 的行列式可以用其各自的行列式的乘积表示

(ix) $$\det(\boldsymbol{AB})=\det(\boldsymbol{BA})=\det\boldsymbol{A}\det\boldsymbol{B} \tag{5}$$

这是行列式的重要性质之一.

[问题 1] 试证明对于可换环 $\tilde{R}$ 上的 $n\times n$ 阶矩阵 $\tilde{R}^{n\times n}$ 用(4)式定义行列式时, 性质(i)～(ix)成立.

设 $\boldsymbol{A}\in F^{n\times n}$ 是正则矩阵, 在(5)式中选取 $\boldsymbol{B}=\boldsymbol{A}^{-1}$, 则

$$\det\boldsymbol{A}\cdot\det\boldsymbol{A}^{-1}=\det(\boldsymbol{AA}^{-1})=\det\boldsymbol{I}=1 \tag{6}$$

由此得

a) 若 $\boldsymbol{A}$ 是正则矩阵,则 $\det\boldsymbol{A}\neq 0$

b) $$\det(\boldsymbol{A}^{-1})=(\det\boldsymbol{A})^{-1} \tag{7}$$

若 $\boldsymbol{A}$ 为正交阵,则 $\det\boldsymbol{A}$ 是 1 或是 -1.

伴随矩阵

考虑以 $\boldsymbol{A}\in F^{n\times n}$ 的元素 a_{ij} 的代数余子式 A_{ij} 为元素的 $n\times n$ 矩阵. 它称为伴随矩阵,用 $\mathrm{adj}\boldsymbol{A}$ 表示,即

$$\mathrm{adj}\boldsymbol{A}\triangleq\begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix} \tag{8}$$

下面来讨论 $\mathrm{adj}\boldsymbol{A}$ 的性质.

首先,由展开式(4)和上述行列式的性质(i),(vii)得

$$\boldsymbol{A}(\mathrm{adj}\boldsymbol{A})=(\mathrm{adj}\boldsymbol{A})\boldsymbol{A}=(\det\boldsymbol{A})\boldsymbol{I}_n \tag{9}$$

现在假定 $\det\boldsymbol{A}\neq 0$, 则由(9)式得

$$\boldsymbol{A}^{-1}=\frac{\mathrm{adj}\boldsymbol{A}}{\det\boldsymbol{A}} \tag{10}$$

即 $\boldsymbol{A}$ 的逆矩阵可以用 $\det\boldsymbol{A}$ 和 $\mathrm{adj}\boldsymbol{A}$ 给出.

由(9),(10)得

c) 对于 $\boldsymbol{A}\in F^{n\times n}$, 若 $\det\boldsymbol{A}\neq 0$, 则 $\boldsymbol{A}^{-1}$ 存在,即 $\boldsymbol{A}$ 是正则矩阵.

d) $$\det(\mathrm{adj}\boldsymbol{A})=(\det\boldsymbol{A})^{n-1} \tag{11}$$

将 a)和 c)合并,可以得到下列关于矩阵正则性的定理.

[定理 1] 矩阵 $\boldsymbol{A}\in F^{n\times n}$ 是正则矩阵的充分和必要条件是 $\det\boldsymbol{A}\neq 0$.

对于在可换环 $\tilde{R}$ 上的 $n\times n$ 阶方阵,用和推导该定理同样的方法可得如下定理.

[定理 2] 矩阵 $\boldsymbol{A}\in\tilde{R}^{n\times n}$ 是幺模阵的充分和必要条件是, $\det\boldsymbol{A}$ 是 $\tilde{R}$ 的可逆元.

定理 1、2 是从 B-0-1 中的例子(2×2 阶方阵)到 $n\times n$ 阶方阵的推广.

[例] 设

$$\boldsymbol{A}\triangleq\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}\in\tilde{R}^{3\times 3}$$

在下列几种情况下

1) $\tilde{R}$ 是整数环 Z

2) $\tilde{R}$ 是实数体 R

3) $\tilde{R}$ 是伽罗瓦体 $GF(2)$

4) $\tilde{R}$ 是伽罗瓦体 $GF(3)$

时, $\boldsymbol{A}^{-1}\in\tilde{R}^{3\times 3}$ 是否存在,若存在的话,求出 $\boldsymbol{A}^{-1}$

1) $\det\boldsymbol{A}=2$, 则不存在.

2)

$$A^{-1}=\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

3) $\det\boldsymbol{A}=0$, 则不存在.

4) $\det\boldsymbol{A}=2\quad(\det\boldsymbol{A})^{-1}=2\quad -2=1$, 则

$$A^{-1}=\begin{bmatrix} 2 & 2 & 1 \\ 1 & 2 & 2 \\ 2 & 1 & 2 \end{bmatrix}$$

拉普拉斯展开式

将行列式对于一个列(由行列式的性质(i)可见, 对于一个行也是一样)展开, 便得到(4)式. 推广到一般情况, 可以对于 r 个列 (或行) 展开, 便得到所谓的拉普拉斯展开式. 为了说明方便,定义下列关于子式的符号.

在 $\boldsymbol{A}\in F^{n\times n}$ 中消去若干行只剩下第 i_1 行、第 i_2 行、…,第 i_r 行 $(1\leqslant i_1<i_2<\cdots<i_r\leqslant n)$, 再消去若干列只剩下第 j_1 列,第 j_2 列, …, 第 j_r 列$(1\leqslant j_1<j_2<\cdots<j_r\leqslant n)$后,便剩下一个 $r\times r$ 阶小方阵, 用符号 $\boldsymbol{A}\begin{pmatrix} i_1, & i_2, & \cdots, & i_r \\ j_1, & j_2, & \cdots, & j_r \end{pmatrix}$ 表示. 这种小矩阵的行列式, 即 $\det\boldsymbol{A}\begin{pmatrix} i_1, & i_2, & \cdots, & i_r \\ j_1, & j_2, & \cdots, & j_r \end{pmatrix}$称为 $\boldsymbol{A}$ 的 r 级子式(minor).

[例] $n=4,\ i_1=1,\ i_2=3,\ j_1=2,\ j_2=3$

$$\boldsymbol{A}=\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}\Rightarrow \boldsymbol{A}\begin{pmatrix} 1 & 3 \\ 2 & 3 \end{pmatrix}=\begin{bmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{bmatrix}$$

设 $\{1,2,\cdots,n\}$ 中除掉 $\{i_1,i_2,\cdots,i_r\}$ 后剩下的是 $\{i'_1,\ i'_2,\cdots,i'_{n-r}\}(i'_1<i'_2<\cdots<i'_{n-r})$. $\{j'_1,\ j'_2,\ \cdots,\ j'_{n-r}\}$ 也同样定义,则

$$\det\boldsymbol{A}=\sum\pm\det\boldsymbol{A}\begin{pmatrix} i_1, & \cdots, & i_r \\ j_1, & \cdots, & j_r \end{pmatrix}\det\boldsymbol{A}\begin{pmatrix} i'_1, & \cdots, & i'_{n-r} \\ j'_1, & \cdots, & j'_{n-r} \end{pmatrix} \tag{12}$$

成立. 式中 $\sum$ 表示对于所有满足 $1\leqslant j_1<j_2<\cdots<j_r\leqslant n$ 的 j 的组合求和. 而且, 当$(i_1+\cdots+i_r+j_1+\cdots+j_r)$为偶数时,符号$\pm$取$+$号,为奇数时取$-$号. (12) 式称为对于第 $i_1,\ i_2,\ \cdots,\ i_r$ 行的拉普拉斯展开式.

[例] 试将 4×4 阶方阵 $\boldsymbol{A}$ 的行列式,对第 1 行和第 3 行展开.

在该例中, $i_1=1,\ i_2=3,\ i'_1=2,\ i'_2=4$. 由 $\{1,\ 2,\ 3,\ 4\}$ 中作出满足 $j_1<j_2$ 的组合, 于是得到 (1, 2), (1, 3), (1, 4), (2, 3), (2, 4), (3, 4) 6 组. 在这里 $\sum$ 变成对这 6 组求和, 即

$$\det A=-\det A\begin{pmatrix}1&3\\1&2\end{pmatrix}\det A\begin{pmatrix}2&4\\3&4\end{pmatrix}+\det A\begin{pmatrix}1&3\\1&3\end{pmatrix}\det A\begin{pmatrix}2&4\\2&4\end{pmatrix}$$
$$-\det A\begin{pmatrix}1&3\\1&4\end{pmatrix}\det A\begin{pmatrix}2&4\\2&3\end{pmatrix}-\det A\begin{pmatrix}1&3\\2&3\end{pmatrix}\det A\begin{pmatrix}2&4\\1&4\end{pmatrix}$$
$$+\det A\begin{pmatrix}1&3\\2&4\end{pmatrix}\det A\begin{pmatrix}2&4\\1&3\end{pmatrix}-\det A\begin{pmatrix}1&3\\3&4\end{pmatrix}\det A\begin{pmatrix}2&4\\1&2\end{pmatrix}$$
$$\det A=-\det\begin{bmatrix}a_{11}&a_{12}\\a_{31}&a_{32}\end{bmatrix}\det\begin{bmatrix}a_{23}&a_{24}\\a_{43}&a_{44}\end{bmatrix}+\det\begin{bmatrix}a_{11}&a_{13}\\a_{31}&a_{33}\end{bmatrix}\det\begin{bmatrix}a_{22}&a_{24}\\a_{42}&a_{44}\end{bmatrix}$$
$$-\det\begin{bmatrix}a_{11}&a_{14}\\a_{31}&a_{34}\end{bmatrix}\det\begin{bmatrix}a_{22}&a_{23}\\a_{42}&a_{43}\end{bmatrix}-\det\begin{bmatrix}a_{12}&a_{13}\\a_{32}&a_{33}\end{bmatrix}\det\begin{bmatrix}a_{21}&a_{24}\\a_{41}&a_{44}\end{bmatrix}$$
$$+\det\begin{bmatrix}a_{12}&a_{14}\\a_{32}&a_{34}\end{bmatrix}\det\begin{bmatrix}a_{21}&a_{23}\\a_{41}&a_{43}\end{bmatrix}-\det\begin{bmatrix}a_{13}&a_{14}\\a_{33}&a_{34}\end{bmatrix}\det\begin{bmatrix}a_{21}&a_{22}\\a_{41}&a_{42}\end{bmatrix}$$

拉普拉斯展开式对于利用 0 元素展开、简化行列式的计算是非常有用的.

设 $A\in F^{n\times n}$, 对于满足 $p+q=n$ 的某些正整数组 p, q 及 i_1, $\cdots$, i_p 和 j_1, $\cdots$, j_q, 若

$$A\begin{pmatrix}i_1, & \cdots, & i_p\\ j_1, & \cdots, & j_q\end{pmatrix}=0_{p,\,n-p} \tag{13}$$

则

$$\det A=\pm\det A\begin{pmatrix}i_1, & \cdots, & i_p\\ j'_1, & \cdots, & j'_p\end{pmatrix}\det A\begin{pmatrix}i'_1, & \cdots, & i'_q\\ j_1, & \cdots, & j_q\end{pmatrix} \tag{14}$$

当$(i_1+\cdots+i_p+j'_1+\cdots+j'_p)$为偶数时式中$\pm$取$+$号,为奇数时取$-$号. 例如,当利用上述展开式展开

$$A=\begin{bmatrix}a_{11}&a_{12}&a_{13}&a_{14}\\a_{21}&0&a_{23}&0\\a_{31}&a_{32}&a_{33}&a_{34}\\a_{41}&0&a_{43}&0\end{bmatrix}$$

时,得

$$\det A=\det A\begin{pmatrix}2&4\\1&3\end{pmatrix}\det A\begin{pmatrix}1&3\\2&4\end{pmatrix}=(a_{21}a_{43}-a_{41}a_{23})(a_{12}a_{34}-a_{32}a_{14})$$

设 $A\in F^{m\times p}$, $B\in F^{p\times n}$, $C\triangleq AB\in F^{m\times n}$,则 C 的子式可以用 A, B 的子式表示[1)]

$$\det C\begin{pmatrix}i_1, & i_2, & \cdots, & i_r\\ j_1, & j_2, & \cdots, & j_r\end{pmatrix}=\sum_k\det A\begin{pmatrix}i_1, & i_2, & \cdots, & i_r\\ k_1, & k_2, & \cdots, & k_r\end{pmatrix}\det B\begin{pmatrix}k_1, & k_2, & \cdots, & k_r\\ j_1, & j_2, & \cdots, & j_r\end{pmatrix}$$
$$r\leqslant\min(n,\ m,\ p) \tag{15}$$

式中 $\sum$ 是对于所有的满足 $1\leqslant k_1<k_2<\cdots<k_r\leqslant p$ 的组合求和.

(12) 式是(4)式的普遍形式. 同样,将(10)式普遍化,可得下面关系.

设 $A\in F^{n\times n}$ 是正则矩阵,则

$$\det A^{-1}\begin{pmatrix}i_1, & i_2, & \cdots, & i_r\\ j_1, & j_2, & \cdots, & j_r\end{pmatrix}=\pm\det A\begin{pmatrix}j'_1, & j'_2, & \cdots, & j'_{n-r}\\ i'_1, & i'_2, & \cdots, & i'_{n-r}\end{pmatrix}/\det A \tag{16}$$

式中$\pm$号按和(10)式同样的原则选取.

---

1) 显然,既使对于长方形矩阵,也可以按着方阵的情况定义子矩阵 $A\begin{pmatrix}i_1, & i_2, & \cdots, & i_r\\ j_1, & j_2, & \cdots, & j_r\end{pmatrix}$.

范德蒙行列式

设 F 是任意体，由 F 中 n 个元 $x_1, x_2, \cdots, x_n$ 组成的体 F 上 $n\times n$ 阶方阵

$$V=\begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ \vdots & \vdots & & \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{bmatrix} \tag{17}$$

的行列式，叫做范德蒙(Vandermonde)行列式．在线性系统的状态方程式中，将系数矩阵化成对角形矩阵时，经常用到这种矩阵．

在范德蒙行列式中，进行第 i 行$-$(第 $(i-1)$ 行的 x_1 倍)$(i=2, 3, \cdots, n)$ 的演算，$\det V$ 的值不变，

$$\det V=\det\begin{bmatrix} 1 & 1 & & 1 \\ 0 & x_2-x_1 & & x_n-x_1 \\ \vdots & \vdots & & \vdots \\ 0 & x_2^{n-2}(x_2-x_1) & & x_n^{n-2}(x_n-x_1) \end{bmatrix}$$

$$=(x_2-x_1)(x_3-x_1)\cdots(x_n-x_1)\det\begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_2 & x_3 & \cdots & x_n \\ \vdots & \vdots & & \vdots \\ x_2^{n-2} & x_3^{n-2} & \cdots & x_n^{n-2} \end{bmatrix}$$

反复进行这种运算便可求得 $\det V$

$$\det V=[(x_2-x_1)\cdots(x_n-x_1)][(x_3-x_2)\cdots(x_n-x_2)]\cdots[(x_n-x_{n-1})]=\prod_{i>j}(x_i-x_j) \tag{18}$$

［问题 2］ 设 $x_1, x_2, \cdots, x_n$ 是复数，试求下列 $n\times n$ 阶复数方阵的行列式．

$$\begin{bmatrix} 1 & 0 & 0 & 1 & \cdots & 1 \\ x_1 & 1 & 0 & x_4 & \cdots & x_n \\ x_1^2 & 2x_1 & 1 & x_4^2 & \cdots & x_n^2 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ x_1^{n-1} & (n-1)x_1^{n-2} & \frac{1}{2}(n-1)(n-2)x_1^{n-3} & x_4^{n-1} & \cdots & x_n^{n-1} \end{bmatrix}$$

$$=\left[\begin{matrix} 1 \\ x_1 \\ \vdots \\ x_1^{n-1} \end{matrix}\quad \frac{d}{dx_1}\begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_1^{n-1} \end{bmatrix}\quad \frac{1}{2}\frac{d^2}{dx_1^2}\begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_1^{n-1} \end{bmatrix}\quad \begin{matrix} 1 & \cdots & 1 \\ x_4 & \cdots & x_n \\ \vdots & & \vdots \\ x_4^{n-1} & \cdots & x_n^{n-1} \end{matrix}\right]\in C^{n\times n} \tag{19}$$

迹

设 $\tilde{R}$ 为任意环，$A=[a_{ij}]$是属于 $\tilde{R}^{n\times n}$ 的矩阵，其主对角线元素 $a_{ii}(i=1, \cdots, n)$之和叫做迹(trace)，用 $\mathrm{tr}A$ 表示，

$$\mathrm{tr}A\triangleq\sum_{i=1}^{n}a_{ii}. \tag{20}$$

设 $^\forall A$, $B\in\tilde{R}^{n\times n}$, $^\forall\alpha\in\tilde{R}$，则关于迹有下列性质：

(i) $\operatorname{tr}(\alpha \boldsymbol{A})=\alpha \operatorname{tr} \boldsymbol{A}$

(ii) $\operatorname{tr}(\boldsymbol{A}+\boldsymbol{B})=\operatorname{tr} \boldsymbol{A}+\operatorname{tr} \boldsymbol{B}$

设 $\widetilde{R}$ 为可换环,因积 $\boldsymbol{AB}$ 是方阵,则对于 $\widetilde{R}$ 上的任意矩阵对 $\boldsymbol{A}$, $\boldsymbol{B}$,

(iii) $\operatorname{tr}(\boldsymbol{AB})=\operatorname{tr}(\boldsymbol{BA})$

成立. 必须注意,式中 $\boldsymbol{A}$, $\boldsymbol{B}$ 本身不必是方阵. 例如,若用实数体上的二次型可得下列关系

(iii)' $\boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x}=\operatorname{tr}(\boldsymbol{x} \boldsymbol{x}^T \boldsymbol{A})=\operatorname{tr}(\boldsymbol{A} \boldsymbol{x} \boldsymbol{x}^T)$.

行列式和迹的微分

行列式、迹是由矩阵确定的纯量. 我们常常想知道,当矩阵的元素稍微发生一些变化时,这些量是如何变化的.

起初先假定矩阵 $\boldsymbol{X}$ 的各元素均完全独立变化. 为了知道对应于元素 x_{ij} 的微小变化 $\det \boldsymbol{X}$ 的变化,先来求 $d\det \boldsymbol{X}/dx_{ij}$.

将 $\det \boldsymbol{X}$ 对第 i 行展开,得

$$\det \boldsymbol{X}=\sum_{k=1}^{n} x_{ik} X_{ik}=x_{ij} X_{ij}+\sum_{\substack{k=1 \\ k \neq j}}^{n} x_{ik} X_{ik} \tag{21}$$

式中 X_{ik} 是元素 x_{ik} 的代数余子式. 因(21)式右边第2项不包含 x_{ij}(试叙述其理由),则

$$\frac{d}{dx_{ij}} \det \boldsymbol{X}=X_{ij}$$

因此,若将以 $d\det \boldsymbol{X}/dx_{ij}$ 为 (i, j) 元素的矩阵用 $d\det \boldsymbol{X}/d\boldsymbol{X}$ 表示,则

$$\frac{d}{d\boldsymbol{X}} \det \boldsymbol{X}=[X_{ij}]=[\operatorname{adj} \boldsymbol{X}]^T=\det \boldsymbol{X}[\boldsymbol{X}^{-1}]^T \tag{22}$$

其次,我们来研究矩阵 $\boldsymbol{X}$ 的所有元素均受一个参数 t 影响时的情况.

求 $d\det \boldsymbol{X}(t)/dt$, 得

$$\frac{d}{dt} \det \boldsymbol{X}(t)=\sum_{i,j=1}^{n} \frac{\partial}{\partial x_{ij}} \det \boldsymbol{X}(t) \frac{dx_{ij}(t)}{dt}=\sum_{i=1}^{n} \sum_{j=1}^{n} X_{ij} \frac{dx_{ij}(t)}{dt}$$

在这里,若考虑只将矩阵 $\boldsymbol{X}$ 的第 i 行对 t 微分得到的矩阵

$$\boldsymbol{X}^{(i)}=\begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{i-1,1} & \cdots & x_{i-1,n} \\ \dot{x}_{i1} & \cdots & \dot{x}_{i,n} \\ x_{i+1,1} & \cdots & x_{i+1,n} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{n,n} \end{bmatrix} \begin{array}{l} \\ \\ \\ \leftarrow \text{第 } i \text{ 行} \\ \\ \\ \\ \end{array} \tag{23}$$

则

$$\sum_{j=1}^{n} X_{ij} \frac{dx_{ij}}{dt}=\sum_{j=1}^{n} \dot{x}_{ij} X_{ij}=\det \boldsymbol{X}^{(i)}$$

由此得

$$\frac{d}{dt} \det \boldsymbol{X}(t)=\sum_{i=1}^{n} \det \boldsymbol{X}^{(i)} \tag{24}$$

对于 $\operatorname{tr} \boldsymbol{X}$ 也一样,在 $\boldsymbol{X}$ 的各元素独立变化的情况下

$$\frac{\partial}{\partial x_{ij}}\operatorname{tr}\boldsymbol{X}=\frac{\partial}{\partial x_{ij}}\sum_{k=1}^{m}x_{kk}=\begin{cases}1 & i=j\\ 0 & i\neq j\end{cases}$$

$$\therefore\quad \frac{\partial\operatorname{tr}\boldsymbol{X}}{\partial\boldsymbol{X}}=\boldsymbol{I}\tag{25}$$

$$\frac{\partial}{\partial x_{ij}}\operatorname{tr}\boldsymbol{AX}=\frac{\partial}{\partial x_{ij}}\sum_{k=1}^{m}\sum_{p}a_{kp}x_{pk}=a_{ji}$$

$$\therefore\quad \frac{\partial\operatorname{tr}\boldsymbol{AX}}{\partial\boldsymbol{X}}=\boldsymbol{A}^T\tag{26}$$

[问题 3]　当 $\boldsymbol{X}$ 是参数 t 的函数 $\boldsymbol{X}(t)$ 时，$d\operatorname{tr}\boldsymbol{X}(t)/dt$ 等于多少？

同理，可得到下列各关系式[21]

$$\frac{\partial}{\partial\boldsymbol{X}}\operatorname{tr}[\boldsymbol{AX}^T]=\boldsymbol{A}\tag{27}$$

$$\frac{\partial}{\partial\boldsymbol{X}}\operatorname{tr}[\boldsymbol{AXB}]=\boldsymbol{A}^T\boldsymbol{B}^T\tag{28}$$

$$\frac{\partial}{\partial\boldsymbol{X}}\operatorname{tr}[\boldsymbol{AX}^T\boldsymbol{B}]=\boldsymbol{BA}\tag{29}$$

$$\frac{\partial}{\partial\boldsymbol{X}^T}\operatorname{tr}[\boldsymbol{AX}]=\boldsymbol{A}\tag{30}$$

$$\frac{\partial}{\partial\boldsymbol{X}^T}\operatorname{tr}[\boldsymbol{AX}^T]=\boldsymbol{A}^T\tag{31}$$

$$\frac{\partial}{\partial\boldsymbol{X}^T}\operatorname{tr}[\boldsymbol{AXB}]=\boldsymbol{BA}\tag{32}$$

$$\frac{\partial}{\partial\boldsymbol{X}^T}\operatorname{tr}[\boldsymbol{AX}^T\boldsymbol{B}]=\boldsymbol{A}^T\boldsymbol{B}^T\tag{33}$$

$$\frac{\partial}{\partial\boldsymbol{X}}\operatorname{tr}[\boldsymbol{XX}]=2\boldsymbol{X}^T\tag{34}$$

$$\frac{\partial}{\partial\boldsymbol{X}}\operatorname{tr}[\boldsymbol{XX}^T]=2\boldsymbol{X}\tag{35}$$

$$\frac{\partial}{\partial\boldsymbol{X}}\operatorname{tr}[\boldsymbol{X}^n]=n(\boldsymbol{X}^{n-1})^T\tag{36}$$

$$\frac{\partial}{\partial\boldsymbol{X}}\operatorname{tr}[\boldsymbol{AX}^n]=\left(\sum_{i=0}^{n-1}\boldsymbol{X}^i\boldsymbol{AX}^{n-1-i}\right)^T\tag{37}$$

$$\frac{\partial}{\partial\boldsymbol{X}}\operatorname{tr}[\boldsymbol{AXBX}]=\boldsymbol{A}^T\boldsymbol{X}^T\boldsymbol{B}^T+\boldsymbol{B}^T\boldsymbol{X}^T\boldsymbol{A}^T\tag{38}$$

$$\frac{\partial}{\partial\boldsymbol{X}}\operatorname{tr}[\boldsymbol{AXBX}^T]=\boldsymbol{A}^T\boldsymbol{X}^T\boldsymbol{B}^T+\boldsymbol{AXB}\tag{39}$$

$$\frac{\partial}{\partial\boldsymbol{X}}\operatorname{tr}[\boldsymbol{X}^{-1}]=-(\boldsymbol{X}^{-1}\boldsymbol{X}^{-1})^T=-(\boldsymbol{X}^{-2})^T\tag{40}$$

$$\frac{\partial}{\partial\boldsymbol{X}}\operatorname{tr}[\boldsymbol{AX}^{-1}\boldsymbol{B}]=-(\boldsymbol{X}^{-1}\boldsymbol{BAX}^{-1})^T\tag{41}$$

$$\frac{\partial}{\partial\boldsymbol{X}}\log\det[\boldsymbol{X}]=(\boldsymbol{X}^{-1})^T\tag{42}$$

$$\frac{\partial}{\partial\boldsymbol{X}}\det[\boldsymbol{X}^T]=\frac{\partial}{\partial\boldsymbol{X}}\det[\boldsymbol{X}]=(\det[\boldsymbol{X}])(\boldsymbol{X}^{-1})^T\tag{43}$$

$$\frac{\partial}{\partial\boldsymbol{X}}\det[\boldsymbol{X}^n]=n(\det[\boldsymbol{X}]^n(\boldsymbol{X}^{-1})^T)\tag{44}$$

B-I-1　分块矩阵的运算、行列式

在 P-I-1 中已讲过矩阵的分块，这里我们再来讲一下分块矩阵的运算问题．在处理较大的矩阵时，除了对原矩阵直接进行运算外，为了减少计算工作量，还要考虑对矩阵进行分块，分块之后不仅运算方便，计算结果表示简洁，而且矩阵之间的相互关系看得更清楚．这里我们只讨论参加运算的矩阵在纵向及横向分割成 2 块或 2 块以下的情况，但是反复使用其结果，很容易推导出在纵向及横向分割成 3 块或更多块情况下的公式．当然，分割方法要满足 P-I-1 中讲过的 (i)、(ii) 两个条件．同时要注意，根据具体情况，可能还要满足其它条件．下面设 $\tilde{R}$ 为任意环．

和

现在假定，将 $\boldsymbol{A}\in\tilde{R}^{m\times n}$，$\boldsymbol{B}\in\tilde{R}^{m\times n}$ 分割成

$$\boldsymbol{A}=\begin{array}{c} \begin{array}{cc} n_1 & n_2 \end{array} \\ \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix} \end{array}\begin{array}{l} m_1 \\ m_2 \end{array} \qquad \boldsymbol{B}=\begin{array}{c} \begin{array}{cc} p_1 & p_2 \end{array} \\ \begin{bmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} \end{bmatrix} \end{array}\begin{array}{l} q_1 \\ q_2 \end{array} \tag{1}$$

除了上述的条件 (i)，(ii) 以外，分割方法还要满足下列条件

$$n_1=p_1,\quad n_2=p_2,\quad m_1=q_1,\quad m_2=q_2 \tag{2}$$

显然

$$\boldsymbol{A}+\boldsymbol{B}=\left[\begin{array}{c|c} \boldsymbol{A}_{11}+\boldsymbol{B}_{11} & \boldsymbol{A}_{12}+\boldsymbol{B}_{12} \\ \boldsymbol{A}_{21}+\boldsymbol{B}_{22} & \boldsymbol{A}_{22}+\boldsymbol{B}_{22} \end{array}\right] \tag{3}$$

积

其次，我们来看 $\boldsymbol{A}\in\tilde{R}^{m\times p}$，$\boldsymbol{B}\in\tilde{R}^{p\times n}$* 的乘积．表 1 中归纳出乘积的普遍公式 a)，以及在其它特殊情况下经常用到的关系 b)～e)．

表 1　分块矩阵的乘积

	$\boldsymbol{A}$	$\boldsymbol{B}$	分割方法	(i)，(ii) 以外的条件	$\boldsymbol{AB}$
a)	$\begin{array}{c}\begin{array}{cc}p_1 & p_2\end{array}\\ \left[\begin{array}{c\|c}\hat{\boldsymbol{A}}_{11} & \hat{\boldsymbol{A}}_{12}\\ \hline \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{array}\right]\end{array}\begin{array}{l}\}m_1\\ \}m_2\end{array}$	$\begin{array}{c}\begin{array}{cc}n_1 & n_2\end{array}\\ \left[\begin{array}{c\|c}\hat{\boldsymbol{B}}_{11} & \hat{\boldsymbol{B}}_{12}\\ \hline \boldsymbol{B}_{21} & \boldsymbol{B}_{22}\end{array}\right]\end{array}\begin{array}{l}\}q_1\\ \}q_2\end{array}$	$\boldsymbol{A}$，$\boldsymbol{B}$ 纵横都分成 2 块	$p_1=q_1$ $p_2=q_2$	$\begin{bmatrix}\boldsymbol{A}_{11}\boldsymbol{B}_{11}+\boldsymbol{A}_{12}\boldsymbol{B}_{21} & \boldsymbol{A}_{11}\boldsymbol{B}_{12}+\boldsymbol{A}_{12}\boldsymbol{B}_{22}\\ \boldsymbol{A}_{21}\boldsymbol{B}_{11}+\boldsymbol{A}_{22}\boldsymbol{B}_{21} & \boldsymbol{A}_{21}\boldsymbol{B}_{12}+\boldsymbol{A}_{22}\boldsymbol{B}_{22}\end{bmatrix}$
b)	$[\boldsymbol{A}]$	$\begin{array}{c}\begin{array}{cc}n_1 & n_2\end{array}\\ [\hat{\boldsymbol{B}}_1 \mid \hat{\boldsymbol{B}}_2]\end{array}$	$\boldsymbol{A}$ 不分割 $\boldsymbol{B}$ 纵向分割	无	$[\boldsymbol{AB}_1 \mid \boldsymbol{AB}_2]$
c)	$\left[\begin{array}{c}\boldsymbol{A}_1\\ \hline \boldsymbol{A}_2\end{array}\right]\begin{array}{l}\}m_1\\ \}m_2\end{array}$	$[\boldsymbol{B}]$	$\boldsymbol{A}$ 横向分割 $\boldsymbol{B}$ 不分割	无	$\left[\begin{array}{c}\boldsymbol{A}_1\boldsymbol{B}\\ \hline \boldsymbol{A}_2\boldsymbol{B}\end{array}\right]$
d)	$\begin{array}{c}\begin{array}{cc}p_1 & p_2\end{array}\\ [\hat{\boldsymbol{A}}_1 \mid \hat{\boldsymbol{A}}_2]\end{array}$	$\left[\begin{array}{c}\boldsymbol{B}_1\\ \hline \boldsymbol{B}_2\end{array}\right]\begin{array}{l}\}q_1\\ \}q_2\end{array}$	$\boldsymbol{A}$ 纵向分割 $\boldsymbol{B}$ 横向分割	$p_1=q_1$ $p_2=q_2$	$[\boldsymbol{A}_1\boldsymbol{B}_1]+[\boldsymbol{A}_2\boldsymbol{B}_2]$
e)	$\left[\begin{array}{c}\boldsymbol{A}_1\\ \hline \boldsymbol{A}_2\end{array}\right]\begin{array}{l}\}m_1\\ \}m_2\end{array}$	$\begin{array}{c}\begin{array}{cc}n_1 & n_2\end{array}\\ [\hat{\boldsymbol{B}}_1 \mid \hat{\boldsymbol{B}}_2]\end{array}$	$\boldsymbol{A}$ 横向分割 $\boldsymbol{B}$ 纵向分割	无	$\left[\begin{array}{c\|c}\boldsymbol{A}_1\boldsymbol{B}_1 & \boldsymbol{A}_1\boldsymbol{B}_2\\ \hline \boldsymbol{A}_2\boldsymbol{B}_1 & \boldsymbol{A}_2\boldsymbol{B}_2\end{array}\right]$

* 原文误为 $\tilde{R}^{q\times n}$．——译者注

[问题 1]　将 $A\in\tilde{R}^{n\times n}$，$B\in\tilde{R}^{n\times n}$ 均分割成下列分块对角形矩阵

$$A=\begin{bmatrix} A_{11} & 0 & \cdots & 0 \\ & & & \vdots \\ 0 & A_{22} & & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & A_{kk} \end{bmatrix},\quad B=\begin{bmatrix} B_{11} & 0 & \cdots & 0 \\ & & & \vdots \\ 0 & B_{22} & & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & B_{kk} \end{bmatrix}$$

其中 A_{ii}，B_{ii} 均为 $n_i\times n_i$ 阶方阵($i=1,\ 2,\ \cdots,\ k,\ n_1+n_2+\cdots+n_k=n$)，试证明

$$AB=\begin{bmatrix} A_{11}B_{11} & 0 & \cdots & 0 \\ 0 & A_{22}B_{22} & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & A_{kk}B_{kk} \end{bmatrix}$$

行列式

对于体 F 上的 $n\times n$ 阶分块三角形矩阵，我们来讨论一下利用其特殊型的行列式的计算方法．设 A 可分割成

$$A=\begin{array}{c} \begin{array}{cc} \;\;n_1 & \;\;n_2 \end{array} \\ \left[\begin{array}{c|c} \overbrace{A_{11}} & \overbrace{A_{12}} \\ \hline 0 & A_{22} \end{array}\right] \end{array}\begin{array}{l} \}n_1 \\ \}n_2 \end{array}\quad n_1+n_2=n \tag{4}$$

式中 $A_{21}=0_{n_2\times n_1}$，利用下列定理求 $\det A$ 是很方便的．

[定理 1]

在(4)式分块三角形矩阵中，

$$\det A=\det A_{11}\cdot\det A_{22} \tag{5}$$

(证明)

由关于第 1 行～第 n_1 行的拉普拉斯展开式得证．

反复使用该定理可得下列结果．

[系 1]

$$\det\begin{bmatrix} \overbrace{A_{11}}^{n_1} & \overbrace{A_{12}}^{n_2} & \cdots & \overbrace{A_{1r}}^{n_r} \\ 0 & A_{22} & \cdots & A_{2r} \\ \vdots & \vdots & & \\ 0 & 0 & \cdots & A_{rr} \end{bmatrix}\begin{array}{l} \}n_1 \\ \}n_2 \\ \vdots \\ \}n_r \end{array}=\det A_{11}\cdot\det A_{22}\cdot\cdots\cdot\det A_{rr} \tag{6}$$

下面再来讨论应用范围更广泛一些的定理．

设 $A\in F^{n\times n}$，F 为任意体，将 A 分割成

$$A=\left[\begin{array}{c|c} \overbrace{A_{11}}^{n_1} & \overbrace{A_{12}}^{n_2} \\ \hline A_{21} & A_{22} \end{array}\right]\begin{array}{l} \}n_1 \\ \}n_2 \end{array}\quad n_1+n_2=n \tag{7}$$

为了求 A 的行列式，这里我们先讨论如何将 A 表示成两个三角形矩阵的乘积．

设 $A_{11}\in F^{n_1\times n_1}$ 是正则矩阵，则

$$\begin{bmatrix} I_{n_1} & 0 \\ -A_{21}A_{11}^{-1} & I_{n_2} \end{bmatrix}\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}=\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22}-A_{21}A_{11}^{-1}A_{12} \end{bmatrix} \tag{8}$$

象这样对 A 左乘以正则矩阵，就意味着对 A 进行下章(B-I-2)所叙述的基本变换．亦即，

(8)式和对 A 进行适当的行变换,变换成(8)式右边的分块上三角形矩阵是一回事.

于是由(8)式得

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I_{n_1} & 0 \\ A_{21}A_{11}^{-1} & I_{n_2} \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22}-A_{21}A_{11}^{-1}A_{12} \end{bmatrix} \tag{9}$$

即 A 表示成两个分块三角形矩阵的乘积.

(9) 式右边第二个矩阵还可以再分解成对角形矩阵和三角形矩阵之积. 若 $\det A_{11} \neq 0$, 则最后可表示成

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I_{n_1} & 0 \\ A_{21}A_{11}^{-1} & I_{n_2} \end{bmatrix} \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22}-A_{21}A_{11}^{-1}A_{12} \end{bmatrix} \begin{bmatrix} I_{n_1} & A_{11}^{-1}A_{12} \\ 0 & I_{n_2} \end{bmatrix} \tag{10}$$

在 (10)式两边取矩阵的行列式,应用行列式的性质(ix)和定理 1,则当 $\det A_{11} \neq 0$ 时,$\det A \neq 0$ 的充分和必要条件是 $\det(A_{22}-A_{21}A_{11}^{-1}A_{12}) \neq 0$.

其次,关于 A 我们再来讨论 A_{22} 是正则矩阵时的情况. 这时和得到(10)式的过程完全一样,可以将 A 分解成

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I_{n_1} & A_{12}A_{22}^{-1} \\ 0 & I_{n_2} \end{bmatrix} \begin{bmatrix} A_{11}-A_{12}A_{22}^{-1}A_{21} & 0 \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} I_{n_1} & 0 \\ A_{22}^{-1}A_{21} & I_{n_2} \end{bmatrix} \tag{11}$$

因此, 当 $\det A_{22} \neq 0$ 时, $\det A \neq 0$ 的充分和必要条件是 $\det(A_{11}-A_{12}A_{22}^{-1}A_{21}) \neq 0$. 归纳以上,得如下定理.

[定理 2]

将 $A \in F^{n\times n}$ 分割成

$$A = \begin{array}{c} \begin{array}{cc} \overbrace{\quad}^{n_1} & \overbrace{\quad}^{n_2} \end{array} \\ \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{array}{l} \}n_1 \\ \}n_2 \end{array} \end{array}, \quad n_1+n_2=n \tag{12}$$

当 $\det A_{11} \neq 0$ 时,则

$$\det A = \det A_{11} \cdot \det(A_{22}-A_{21}A_{11}^{-1}A_{12}) \tag{13}$$

成立.

若 $\det A_{22} \neq 0$, 则

$$\det A = \det A_{22} \cdot \det(A_{11}-A_{12}A_{22}^{-1}A_{21}). \tag{14}$$

下面系 2 是定理 2 的特殊情况,在系统理论上经常用到.

[系 2]

设 $A \in F^{m\times n}$, $B \in F^{n\times m}$, 则

$$\det(I_m+AB) = \det(I_n+BA) \tag{15}$$

特别是,对于 $a \in F^{1\times n}$, $b \in F^{n\times 1}$, 有

$$\det[I_n+ba] = 1+ab \tag{16}$$

(证明)

由定理 2 得

$$\det \begin{bmatrix} I_n & B \\ -A & I_m \end{bmatrix} = \det(I_m+AB) = \det(I_n+BA) \tag{17}$$

(问题 2) 分块行列式的计算虽很方便,但不能无条件地使用. 一般将 $A \in F^{2n\times 2n}$ 分割成

$$A=\left[\begin{array}{c|c}\overbrace{A_{11}}^{n} & \overbrace{A_{12}}^{n} \\ \hline A_{21} & A_{22}\end{array}\right]\begin{array}{l}\}n\\ \}n\end{array} \tag{18}$$

时，试证明下式是否成立：

$$\det\boldsymbol{A}=\det\boldsymbol{A}_{11}\cdot\det\boldsymbol{A}_{22}-\det\boldsymbol{A}_{12}\cdot\det\boldsymbol{A}_{21} \tag{19}$$

（问题 3） 对于问题 2 中的 $\boldsymbol{A}$，试证明

$$\det\boldsymbol{A}=\det(\boldsymbol{A}_{11}\boldsymbol{A}_{22}-\boldsymbol{A}_{11}\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}) \quad (\text{若 } \boldsymbol{A}_{11} \text{ 是正则矩阵}) \tag{20a}$$

$$=\det(\boldsymbol{A}_{22}\boldsymbol{A}_{11}-\boldsymbol{A}_{22}\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}) \quad (\text{若 } \boldsymbol{A}_{22} \text{ 是正则矩阵}) \tag{20b}$$

若 $\boldsymbol{A}_{11}$ 和 $\boldsymbol{A}_{21}$（$\boldsymbol{A}_{22}$ 和$\boldsymbol{A}_{12}$）的积可以交换，则

$$\det\boldsymbol{A}=\det(\boldsymbol{A}_{11}\boldsymbol{A}_{22}-\boldsymbol{A}_{21}\boldsymbol{A}_{12}) \tag{21a}$$

$$(\det\boldsymbol{A}=\det(\boldsymbol{A}_{22}\boldsymbol{A}_{11}-\boldsymbol{A}_{12}\boldsymbol{A}_{21})) \tag{21b}$$

（问题 4） 对于问题 2 中的 $\boldsymbol{A}$，设 $\boldsymbol{A}_{11}$ 和 $\boldsymbol{A}_{22}$ 都不是正则矩阵，而 $\boldsymbol{A}_{12}$（或 $\boldsymbol{A}_{21}$）是正则矩阵，试证明 $\det\boldsymbol{A}$ 可表示成什么形式？

（问题 5） 对于问题 2 中的 $\boldsymbol{A}$，当

$$\boldsymbol{A}_{11}+\boldsymbol{A}_{21}=\boldsymbol{A}_{12}+\boldsymbol{A}_{22} \tag{22}$$

时，试证明

$$\det\boldsymbol{A}=\det(\boldsymbol{A}_{11}+\boldsymbol{A}_{21})\det(\boldsymbol{A}_{21}-\boldsymbol{A}_{22}) \tag{23}$$

逆矩阵

下面我们来讨论 $\boldsymbol{A}$ 分块后的逆矩阵的计算方法．我们知道，$\boldsymbol{A}$ 分块后可以分解成(10)，(11)式所示的三个矩阵的乘积．因此，为了求 $\boldsymbol{A}^{-1}$，可先求出上述三个矩阵各自的逆矩阵，然后依相反顺序连乘求得，即由(10)式得

$$\begin{aligned}&\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{bmatrix}^{-1}\\ &=\begin{bmatrix}\boldsymbol{I}_{n_1} & -\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}\\ \boldsymbol{0} & \boldsymbol{I}_{n_2}\end{bmatrix}\begin{bmatrix}\boldsymbol{A}_{11}^{-1} & \boldsymbol{0}\\ \boldsymbol{0} & (\boldsymbol{A}_{22}-\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12})^{-1}\end{bmatrix}\begin{bmatrix}\boldsymbol{I}_{n_1} & \boldsymbol{0}\\ -\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1} & \boldsymbol{I}_{n_2}\end{bmatrix}\end{aligned} \tag{24}$$

另外，由(11)式得

$$\begin{aligned}&\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22}\end{bmatrix}^{-1}\\ &=\begin{bmatrix}\boldsymbol{I}_{n_1} & \boldsymbol{0}\\ -\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1} & \boldsymbol{I}_{n_2}\end{bmatrix}\begin{bmatrix}(\boldsymbol{A}_{11}-\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21})^{-1} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{A}_{22}^{-1}\end{bmatrix}\begin{bmatrix}\boldsymbol{I}_{n_1} & -\boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\\ \boldsymbol{0} & \boldsymbol{I}_{n_2}\end{bmatrix}\end{aligned} \tag{25}$$

归纳上述乘式得出如下定理．

［定理 3］ 关于(7)式所示分块矩阵有下列性质．

(i) 若 $\boldsymbol{A}_{11}$ 是正则矩阵，则

$$\boldsymbol{A} \text{ 是正则矩阵} \Leftrightarrow (\boldsymbol{A}_{22}-\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}) \text{ 是正则矩阵}$$

(ii) 若 $\boldsymbol{A}_{11}$，$(\boldsymbol{A}_{22}-\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12})$ 是正则矩阵，由(i)可知 $\boldsymbol{A}^{-1}$ 存在．

$$\boldsymbol{A}^{-1} \text{ 的}(1,\ 1)\text{子块}=\boldsymbol{A}_{11}^{-1}+\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}(\boldsymbol{A}_{22}-\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12})^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1} \tag{26a}$$

$$\boldsymbol{A}^{-1} \text{ 的}(1,\ 2)\text{子块}=-\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}(\boldsymbol{A}_{22}-\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12})^{-1} \tag{26b}$$

$$\boldsymbol{A}^{-1} \text{ 的}(2,\ 1)\text{子块}=-(\boldsymbol{A}_{22}-\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12})^{-1}\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1} \tag{26c}$$

$$\boldsymbol{A}^{-1} \text{ 的}(2,\ 2)\text{子块}=(\boldsymbol{A}_{22}-\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12})^{-1} \tag{26d}$$

(iii) 若 A_{22} 是正则矩阵，则

$$A \text{ 是正则矩阵} \Leftrightarrow (A_{11}-A_{12}A_{22}^{-1}A_{21}) \text{ 是正则矩阵}$$

(iv) 若 A_{22}，$(A_{11}-A_{12}A_{22}^{-1}A_{21})$ 是正则矩阵，由(iii)可知，A^{-1} 存在.

$$A^{-1} \text{ 的 } (1,\ 1) \text{ 子块} = (A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1} \tag{27a}$$

$$A^{-1} \text{ 的 } (1,\ 2) \text{ 子块} = -(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}A_{12}A_{22}^{-1} \tag{27b}$$

$$A^{-1} \text{ 的 } (2,\ 1) \text{ 子块} = -A_{22}^{-1}A_{21}(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1} \tag{27c}$$

$$A^{-1} \text{ 的 } (2,\ 2) \text{ 子块} = A_{22}^{-1}+A_{22}^{-1}A_{21}(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}A_{12}A_{22}^{-1} \tag{27d}$$

(v) 若 A_{11}，A_{22} 均为正则矩阵，当 $(A_{22}-A_{21}A_{11}^{-1}A_{12})$ 和 $(A_{11}-A_{12}A_{22}^{-1}A_{21})$ 中有一个是正则矩阵时，则另外一个及 A 也是正则矩阵.

(vi) 若 A_{11}，A_{22} 均为正则矩阵，而且 A 也是正则矩阵，则由(26)，(27)式得到下列恒等式

$$(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}=A_{11}^{-1}+A_{11}^{-1}A_{12}(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1} \tag{28a}$$

$$(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}A_{12}A_{22}^{-1}=A_{11}^{-1}A_{12}(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1} \tag{28b}$$

$$A_{22}^{-1}A_{21}(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}=(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1} \tag{28c}$$

$$A_{22}^{-1}+A_{22}^{-1}A_{21}(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}A_{12}A_{22}^{-1}=(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1} \tag{28d}$$

[系 3] 若 $A_{21}=0$，则

(i) A_{11} 和 A_{22} 均为正则矩阵 $\Leftrightarrow A$ 为正则矩阵

(ii)
$$\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{bmatrix} \tag{29}$$

B-I 2 矩阵的基本变换

本章讨论矩阵的基本行变换．它使矩阵变成尽可能简单的形式而它不改变矩阵的基本性质．这在研究线性方程组的解及在原讲座计划写的第 II 部分线性变换的性质时具有重要意义．

基本行(列)变换

对矩阵 $\boldsymbol{A}\in F^{m\times n}$ 进行的下列三种类型的变换都叫做基本行变换 (elementary row operation)．

(第一种类型) 任意一行(例如第 i 行，$1\leqslant i\leqslant m$) 乘以不为零的 $\alpha\in F$．显然，这和 $\boldsymbol{A}$ 左乘以矩阵 $\boldsymbol{T}_{r1}$ 是等价的．

$$\boldsymbol{A}\xrightarrow{\text{将}\boldsymbol{A}\text{的第}i\text{行乘以}\alpha\neq 0,}\boldsymbol{A}_{r1}=\boldsymbol{T}_{r1}\boldsymbol{A}$$

式中

$$\boldsymbol{T}_{r1}=\begin{array}{c}\begin{bmatrix}1 & & & & & & \\ & \ddots & & & & 0 & \\ & & 1 & & & & \\ & & & \alpha & & & \\ & & & & 1 & & \\ & 0 & & & & \ddots & \\ & & & & & & 1\end{bmatrix}\\ \uparrow\ \text{第 } i \text{ 列}\end{array}\begin{array}{l}\\ \\ \\ \leftarrow\text{第 } i\text{ 行}\\ \\ \\ \\ \\ \end{array}=\begin{bmatrix}\boldsymbol{e}_1^T\\ \vdots\\ \boldsymbol{e}_{i-1}^T\\ \alpha\boldsymbol{e}_i^T\\ \boldsymbol{e}_{i+1}^T\\ \vdots\\ \boldsymbol{e}_n^T\end{bmatrix}\tag{1}$$

$\boldsymbol{T}_{r1}$ 是在单位矩阵 $\boldsymbol{I}_m$ 的第 i 行乘以 α 得到的矩阵，即 $\boldsymbol{T}_{r1}$ 是在 $\boldsymbol{I}_m$ 上进行第一种类型基本变换得到的．

(第二种类型) 任意一行(第 i 行，$1\leqslant i\leqslant m$)加上其它一行(第 j 行，$1\leqslant j\leqslant m, i\neq j$)乘以任意的 $\alpha\in F$．这种变换和将矩阵 $\boldsymbol{A}$ 左乘以 $m\times m$ 方阵 $\boldsymbol{T}_{r2}$ 是一样的．

$$\boldsymbol{A}\xrightarrow{\text{将}\boldsymbol{A}\text{的第}j\text{行乘以}\alpha\text{后加到第}i\text{行,}}\boldsymbol{A}_{r2}=\boldsymbol{T}_{r2}\boldsymbol{A}$$

$$\boldsymbol{T}_{r2}=\begin{array}{c}\begin{bmatrix}1 & & & & & 0 & \\ & \ddots & & & & & \\ & & 1 & 0\cdots 0 & & \alpha & \\ & & & \ddots & & 0 & \\ & & & & \ddots & \vdots & \\ & & & & & 0 & \\ & & & & & 1 & \\ & & & & & & \ddots \\ & & & & & & & 1\end{bmatrix}\\ \text{第 } i \text{ 列}\qquad\text{第 } j \text{ 列}\end{array}\begin{array}{l}\\ \\ \leftarrow\text{第 } i\text{ 行}\\ \\ \\ \\ \leftarrow\text{第 } j\text{ 行}\\ \\ \\ \end{array}=\begin{bmatrix}\boldsymbol{e}_1^T\\ \vdots\\ \boldsymbol{e}_{i-1}^T\\ \boldsymbol{e}_i^T+\alpha\boldsymbol{e}_j^T\\ \boldsymbol{e}_{i+1}^T\\ \vdots\\ \boldsymbol{e}_n^T\end{bmatrix}\tag{2}$$

$\boldsymbol{T}_{r2}$ 是将 $\boldsymbol{I}_m$ 的第 j 行乘上 α 加到第 i 行后得到的．

（第三种类型） 两个行互换

第 i 行和第 j 行互换可以由 $\boldsymbol{A}$ 左乘以 $m\times m$ 阶方阵 $\boldsymbol{T}_{r3}$ 表示.

$$\boldsymbol{A}\xrightarrow{\text{将}\boldsymbol{A}\text{的第}i\text{行和第}j\text{行互换}}\boldsymbol{A}_{r3}=\boldsymbol{T}_{r3}\boldsymbol{A}$$

$$\boldsymbol{T}_{r3}=\begin{bmatrix}1&\cdots&&&&&&&0\\&\ddots&&&&&&&\\&&1&&&&&&\\&&&0&\cdots&0&1&&\\&&&\vdots&1&&0&&\\&&&0&&1&\vdots&&\\&&&1&0&\cdots&0&&\\&&&&&&&1&\\&&&&&&&&\ddots\\0&\cdots&&&&&&&1\end{bmatrix}\begin{matrix}\\\\\\\leftarrow\text{第 }i\text{ 行}\\\\\\\leftarrow\text{第 }j\text{ 行}\\\\\\\\\end{matrix}=\begin{bmatrix}\boldsymbol{e}_1^T\\\vdots\\\boldsymbol{e}_{i-1}^T\\\boldsymbol{e}_i^T\\\boldsymbol{e}_{i+1}^T\\\vdots\\\boldsymbol{e}_{j-1}^T\\\boldsymbol{e}_j^T\\\boldsymbol{e}_{j+1}^T\\\vdots\\\boldsymbol{e}_n^T\end{bmatrix}\qquad(3)$$

$$\qquad\downarrow\text{第 }i\text{ 列}\qquad\downarrow\text{第 }j\text{ 列}$$

$\boldsymbol{T}_{r3}$ 是在 $\boldsymbol{I}_m$ 上将第 i 行和第 j 行互换后得到的. $\boldsymbol{T}_{r3}$ 是顺序矩阵.

（例） 现举个对 $3\times n$ 阶方阵进行基本行变换的例子. 设 $\boldsymbol{A}$ 的第 1, 2, 3 行分别为 $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$,

第一种类型: $\begin{bmatrix}1&0&0\\0&\alpha&0\\0&0&1\end{bmatrix}\begin{bmatrix}\boldsymbol{a}_1\\\boldsymbol{a}_2\\\boldsymbol{a}_3\end{bmatrix}=\begin{bmatrix}\boldsymbol{a}_1\\\alpha\boldsymbol{a}_2\\\boldsymbol{a}_3\end{bmatrix}$ 第 2 行乘上 α

第二种类型: $\begin{bmatrix}1&0&0\\0&1&0\\\alpha&0&1\end{bmatrix}\begin{bmatrix}\boldsymbol{a}_1\\\boldsymbol{a}_2\\\boldsymbol{a}_3\end{bmatrix}=\begin{bmatrix}\boldsymbol{a}_1\\\boldsymbol{a}_2\\\alpha\boldsymbol{a}_1+\boldsymbol{a}_3\end{bmatrix}$ 第 1 行乘上 α 加到第 3 行

第三种类型: $\begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}\begin{bmatrix}\boldsymbol{a}_1\\\boldsymbol{a}_2\\\boldsymbol{a}_3\end{bmatrix}=\begin{bmatrix}\boldsymbol{a}_3\\\boldsymbol{a}_2\\\boldsymbol{a}_1\end{bmatrix}$ 第 1 行和第 3 行互换

和基本行变换一样，下列三种变换叫做对于 $\boldsymbol{A}\in F^{m\times n}$ 的基本列变换 (elementary column operation).

（第一种类型） 任意一列乘上 $\alpha\in F(\alpha\neq0)$.

（第二种类型） 任意一列乘上 $\alpha(\alpha\in F)$ 后加到其它一列.

（第三种类型） 两个列互换.

和关于行的情况一样，这种变换和将 $\boldsymbol{A}$ 右乘以下面的 $n\times n$ 阶方阵是一样的，即

（第一种类型）

$$\boldsymbol{A}\xrightarrow{\boldsymbol{A}\text{ 的第 }i\text{ 列乘上 }\alpha\neq0}\boldsymbol{A}_{c1}=\boldsymbol{A}\boldsymbol{T}_{c1}$$

（第二种类型）

$$\boldsymbol{A}\xrightarrow{\boldsymbol{A}\text{ 的第 }j\text{ 列乘上 }\alpha\text{ 后加到第 }i\text{ 列}}\boldsymbol{A}_{c2}=\boldsymbol{A}\boldsymbol{T}_{c2}$$

（第三种类型）

$$\boldsymbol{A}\xrightarrow{\boldsymbol{A}\text{ 的第 }i\text{ 列和第 }j\text{ 列互换}}\boldsymbol{A}_{c3}=\boldsymbol{A}\boldsymbol{T}_{c3}$$

式中 $\boldsymbol{T}_{c1}(\boldsymbol{T}_{c2}, \boldsymbol{T}_{c3})$ 是单位矩阵 $\boldsymbol{I}_n$ 进行第一种类型（第二种类型、第三种类型）基本列变换后得到的.

（问题 1） 试证明基本行(列)变换矩阵都是正则矩阵. 其次，证明基本行（列）变换矩阵 $\boldsymbol{T}_{ri}(\boldsymbol{T}_{ci})$ $(i=1, 2, 3)$ 的逆矩阵仍然是同一类型的基本行(列)变换矩阵.

（问题 2） 试证明第三种类型的基本行（列）变换是将第一种类型、第二种类型的基本行(列)变换组合后得到的.

对 $\boldsymbol{A}\in F^{m\times n}$ 进行有限次基本行变换后得到的 $\tilde{\boldsymbol{A}}\in F^{m\times n}$，通常称为和 $\boldsymbol{A}$ 行等价(row equivalent)，表示成 $\tilde{\boldsymbol{A}}\overset{r}{\sim}\boldsymbol{A}$. 若 $\tilde{\boldsymbol{A}}\overset{r}{\sim}\boldsymbol{A}$，则可写成

$$\tilde{\boldsymbol{A}}=\boldsymbol{T}_r\boldsymbol{A} \tag{4}$$

式中 $\boldsymbol{T}_r$ 是由有限个基本变换矩阵的乘积构成的 $m\times m$ 正则矩阵.

（问题 3） 对于行等价，试证明下列性质成立.

(i) $\boldsymbol{A}\overset{r}{\sim}\boldsymbol{A}$

(ii) $\boldsymbol{A}\overset{r}{\sim}\boldsymbol{B}\Rightarrow\boldsymbol{B}\overset{r}{\sim}\boldsymbol{A}$

(iii) $\boldsymbol{A}\overset{r}{\sim}\boldsymbol{B},\ \boldsymbol{B}\overset{r}{\sim}\boldsymbol{C}\Rightarrow\boldsymbol{A}\overset{r}{\sim}\boldsymbol{C}$

（问题 4） 试按与上述同样的方法定义列等价，并叙述其性质.

行(列)标准形和行(列)秩

现在我们来讨论如何利用有限次基本行变换，把任意矩阵 $\boldsymbol{A}\in F^{m\times n}$ 变换成应用方便的阶梯形矩阵. 最后将得到这样的结论：利用基本行变换可以将 $\boldsymbol{A}\in F^{m\times n}$ 变换成下列形式的 $m\times n$ 矩阵

$$\tilde{\boldsymbol{A}}_r=\begin{array}{c} \begin{array}{ccccccc} & \text{第 } k_1 \text{ 列} & & \text{第 } k_2 \text{ 列} & & \text{第 } k_r \text{ 列} & \\ & \downarrow & & \downarrow & & \downarrow & \end{array} \\ \left[\begin{array}{ccccccc} 0\cdots0 & 1 & *\cdots* & 0 & *\cdots* & 0 & *\cdots* \\ 0\cdots0 & 0 & 0\cdots0 & 1 & *\cdots* & 0 & *\cdots* \\ & & & & & \vdots & \vdots \\ & & & & & 0 & \vdots \\ 0\cdots & & & & \cdots0 & 1 & *\cdots* \\ 0\cdots & & & & & & \cdots0 \\ \vdots & & & & & & \vdots \\ 0\cdots & & & & & & \cdots0 \end{array}\right] \begin{array}{l} \left.\begin{array}{l} \\ \\ \\ \\ \\ \end{array}\right\} r \text{ 个不全为 0 的行} \\ \left.\begin{array}{l} \\ \\ \\ \end{array}\right\} (m-r) \text{ 个全为 0 的行} \end{array} \end{array} \tag{5}$$

* 表示不一定是 0 的元素.

在上面的简化矩阵中，对于某些整数 $r(0\leqslant r\leqslant m)$ 和 $k_1, k_2, \cdots, k_r(1\leqslant k_i\leqslant n)$，下列条件成立.

(i) 第 $1, 2, \cdots, r$ 行左数第一个等于 1(而不等于 0)的元素分别在第 $k_1, k_2, \cdots, k_r$ 列出现.

(ii) $k_1<k_2<\cdots<k_r$，即第 $(i+1)$ 的 1 行与第 i 行相比，前者第一个等于 1 的元素出现在右边. 因此，$i=1, 2, \cdots, r-1$.

(iii) 第 $(r+1)$ 行，…，第 m 行的元素全为 0. 因此，$0\leqslant r\leqslant m$.

(iv) 在第 k_i 列中，除第 i 行的等于 1 的元素外，其它元素全为 0.

(例)

$$\begin{bmatrix} 1 & 4 & 0 & 7 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

是 $r=2, k_1=1, k_2=3$ 的简化矩阵. 因此, 为了利用基本行变换将任意矩阵 $\boldsymbol{A}\in F^{m\times n}$ 变换成(5)式的形式,例如可采用下列方法.

[求 $\tilde{\boldsymbol{A}}_r$ 的方法]

设所给的 $m\times n$ 矩阵为

$$\boldsymbol{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}\in F^{m\times n} \tag{6}$$

(i) 在第 1 行,从左起找出第一个不为 0 的元素,设其为位于第 j_1 列的元素 a_{1j_1}. 用适当的数乘第 1 行,然后将第 1 行加到第 2 行,使 $a_{2j_1}=0$(第二种类型的初等变换). 以下再进行同样的变换(第二种类型),使第 j_1 列位于 a_{1j_1} 元素以下的元素全变成 0, 即 $(2, j_1)$ 元素 $=(3, j_1)$ 元素 $=\cdots=(m, j_1)$ 元素 $=0$.

(ii) 其次, 在第 2 行再从左起找出第一个不为 0 的元素, 设其为 $(2, j_2)$, 再利用第二种类型的基本行变换,使第 j_2 列位于 $(2, j_2)$ 元素以下的元素全变为 0.

(iii) 以下进行同样的变换,一直到第 $(m-1)$ 行为止.

(iv) 适当地将两个行互换(第三种类型),使之变成满足条件(ii)的阶梯形.

(v) 对于(iv)中最后得到的阶梯形矩阵,在第 1 行～第 r 行中,从各行左边起,用其第一个不为 0 的元素的倒数分别乘各行(第一种类型).

(vi) 利用第二种类型的初等变换. 使之满足条件(iv).

利用基本行变换将 $\boldsymbol{A}$ 变成 (5) 式形式不限于上述一个方法. 例如, 可以先将 $\boldsymbol{A}$ 的第 1 行和第 2 行互换, 之后再进行上述(i)～(vi)的运算. 此外还可以想出很多不同的方法. 对此,下列结果是很重要的.

[定理 1] 设 $\tilde{\boldsymbol{A}}_r$ 是 $\boldsymbol{A}\in F^{m\times n}$ 的(5)式形式的行等价矩阵. 则 $\tilde{\boldsymbol{A}}_r$ 是唯一的, 与基本行变换的顺序等无关.

(证明方法) 设 $\boldsymbol{A}\in F^{m\times n}$ 可变换成 (5) 式形式的矩阵 $\tilde{\boldsymbol{A}}_{r1}$ 和 $\tilde{\boldsymbol{A}}_{r2}$. 因 $\tilde{\boldsymbol{A}}_{r1}$ 和 $\tilde{\boldsymbol{A}}_{r2}$ 皆与 $\boldsymbol{A}$ 行等价,由问题 3 的性质(ii), (iii)得知, $\tilde{\boldsymbol{A}}_{r1}$ 和 $\tilde{\boldsymbol{A}}_{r2}$ 互为行等价. 因此,根据(4)式,必有满足

$$\tilde{\boldsymbol{A}}_{r2}=\boldsymbol{P}_1\tilde{\boldsymbol{A}}_{r1}$$
$$\tilde{\boldsymbol{A}}_{r1}=\boldsymbol{P}_2\tilde{\boldsymbol{A}}_{r2}$$

的正则矩阵 $\boldsymbol{P}_1$, $\boldsymbol{P}_2$ 存在. 比较上式两边,得 $\boldsymbol{P}_1=\boldsymbol{P}_2=\boldsymbol{I}$, 即 $\tilde{\boldsymbol{A}}_{r1}=\tilde{\boldsymbol{A}}_{r2}$.

(问题 5) 试完成上述证明.

(例)

$$\boldsymbol{A}=\begin{bmatrix} 1 & 3 & 2 & 1 \\ 2 & 0 & 1 & 2 \\ 3 & 1 & 2 & 3 \end{bmatrix}\xrightarrow{\text{第(i)步}}\begin{bmatrix} 1 & 3 & 2 & 1 \\ 0 & -6 & -3 & 0 \\ 0 & -8 & -4 & 0 \end{bmatrix}\xrightarrow{\text{第(ii), (iii), (iv)步}}\begin{bmatrix} 1 & 3 & 2 & 1 \\ 0 & -6 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow{\text{第(v)步}}\begin{bmatrix} 1 & 3 & 2 & 1 \\ 0 & 1 & 1/2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\xrightarrow{\text{第(vi)步}}\begin{bmatrix} 1 & 0 & 1/2 & 1 \\ 0 & 1 & 1/2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}=\tilde{\boldsymbol{A}}_r$$

（问题 6） 注意在上述步骤 (iv)～(vi) 中，不改变元素不全为 0 的行的数目 r，若仅为了知道 r，做到第 (iii) 步已经足够了，试加以证明.

根据定理 1，因为 (5) 式形式的矩阵是唯一的，与基本行变换的次序无关，所以称它为 $\boldsymbol{A}$ 的行标准形. 另外，在行标准形中，元素不全为 0 的行数 r 称为行秩. 由定理 1 很容易得出下列系.

［系 1］ 所有行等价矩阵都具有同一个行标准形.

［系 2］ 进行基本行变换，行秩 r 不变.

［系 3］ 任意正则矩阵 $\boldsymbol{A}\in F^{n\times n}$ 与单位矩阵行等价.

［系 4］ 任意正则矩阵 $\boldsymbol{A}\in F^{n\times n}$ 可以表示成有限个基本行变换矩阵之积.

（问题 7） 试证明系 1，2，3，4.

另外，和上述行变换完全一样. 显然，有限次基本列变换也可以将任意矩阵 $\boldsymbol{A}\in F^{m\times n}$ 变换成下列阶梯形矩阵

$$
\widetilde{\boldsymbol{A}}_c=\left[\begin{array}{ccccccc}
0 & 0 & \cdots & 0 & 0 & \cdots & 0\\
\vdots & \vdots & & \vdots & \vdots & & \vdots\\
0 & \vdots & & & & & \\
1 & \vdots & & & & & \\
* & \vdots & & & & & \\
\vdots & \vdots & & & & & \\
* & 0 & & & & & \\
0 & 1 & & & & & \\
* & * & & & & & \\
\vdots & \vdots & & & & & \\
* & * & & & & & \\
0 & 0\cdots 0 & & 1 & \vdots & & \vdots\\
* & * & & * & \vdots & & \vdots\\
\vdots & \vdots & & \vdots & \vdots & & \vdots\\
* & * & & * & 0 & \cdots & 0
\end{array}\right]
\begin{array}{l}
\\
\\
\\
\leftarrow \text{第 } l_1 \text{ 行}\\
\\
\\
\\
\leftarrow \text{第 } l_2 \text{ 行}\\
\\
\\
\\
\leftarrow \text{第 } l_r \text{ 行}\\
\\
\\
\\
\end{array}
\tag{7}
$$

r 个不全为 0 的列　　$(n-r)$ 个全为 0 的列

* 表示不一定为 0 的元素.

在 $\widetilde{\boldsymbol{A}}_c$ 中，对于某些整数 $r\,(0\leqslant r\leqslant n)$ 和 $l_1,\ l_2,\ \cdots,\ l_r\,(1\leqslant l_i\leqslant m)$，存在下列条件：

(i) 在第 1，2，…，r 列，从上数第一个不为 0 而等于 1 的元素分别位于第 l_1 行、第 l_2 行，…，第 l_r 行.

(ii) $l_1<l_2<\cdots<l_r$，即在前 r 个列上位于各列上面等于 0 的元素的个数，从左到右顺次增加.

(iii) 第 $(r+1)$，…，n 列上的元素全为 0.

(iv) 第 l_i 行上除了第 i 列上的元素等于 1 以外，其它所有的元素全为 0.

［求 $\widetilde{\boldsymbol{A}}_c$ 的方法］

(i) 在第 1 列上，从上面数起，找出第一个不为 0 的元素，设为 $a_{j_1 1}$，将第 1 列乘上

适当的数加到第 2 列，使 $a_{j_12}=0$(第二种类型变换). 以下进行同样变换(第二种类型)，使第 j_1 行上位于 a_{j_11} 元素右边的元素全变成 0, 即 $(j_1, 2)$ 元素 $=(j_1, 3)$ 元素 $=\cdots$, (j_1, n) 元素 $=0$.

(ii) 其次，在第 2 列上也找出第一个不为 0 的元素，设其为 $(j_2, 2)$. 利用第二种类型列变换使位于 j_2 行 $(j_2, 2)$ 元素右边的元素全变成 0.

(iii) 以下进行同样变换，到第 $(n-1)$ 列为止.

(iv) 适当地将两列互换(第三种类型列变换)，使之满足条件(ii).

(v) 对于(iv)中最后得到的阶梯形矩阵，在第 1 列～第 r 列中，各列从上边起，用其第一个不为 0 的元素的倒数分别乘各列(第一种类型).

(vi) 利用第二种类型的基本列变换使之满足条件(iv).

(例)

$$\boldsymbol{A}=\begin{bmatrix}1&3&2&1\\2&0&1&2\\3&1&2&3\end{bmatrix}\xrightarrow{\text{第(i)步}}\begin{bmatrix}1&0&0&0\\2&-6&-3&0\\3&-8&-4&0\end{bmatrix}\xrightarrow{\text{第(ii), (iii), (iv)步}}\begin{bmatrix}1&0&0&0\\2&-6&0&0\\3&-8&-0&0\end{bmatrix}$$

$$\xrightarrow{\text{第(v)步}}\begin{bmatrix}1&0&0&0\\2&1&0&0\\3&4/3&0&0\end{bmatrix}\xrightarrow{\text{第(vi)步}}\begin{bmatrix}1&0&0&0\\0&1&0&0\\1/3&4/3&0&0\end{bmatrix}=\widetilde{\boldsymbol{A}}_c$$

关于列变换也可以得出和定理 1 同样的结论. 因此，将(7)式形式的矩阵称为列标准形，元素不全为 0 的列数叫做列秩. 同样，也可以得出对应于系 1, 2, 3, 4 的结论.

将 $\boldsymbol{A}\in F^{m\times n}$ 的行标准形 $\widetilde{\boldsymbol{A}}_r$ 再进行有限次基本列变换，可以变成下列形式

$$\widetilde{\boldsymbol{A}}\triangleq\left[\begin{array}{c|c}\boldsymbol{I}_r&\boldsymbol{0}_{r,\,n-r}\\\hline\boldsymbol{0}_{m-r,\,r}&\boldsymbol{0}_{m-r,\,n-r}\end{array}\right]\tag{8}$$

(问题 8) 试证明上述结论.

$\widetilde{\boldsymbol{A}}$ 叫做矩阵 $\boldsymbol{A}$ 的标准形，是唯一确定的，与进行行变换、列变换的顺序无关.

基本变换的一个实际应用

最后，作为基本变换的简单应用，介绍一下求正则矩阵 $\boldsymbol{A}\in F^{n\times n}$ 的逆矩阵的方法. 将单位矩阵 $\boldsymbol{I}_n$ 和 $\boldsymbol{A}$ 组合起来，得到下列 $n\times 2n$ 阶矩阵

$$\boldsymbol{B}=[\boldsymbol{A}\,|\,\boldsymbol{I}_n]\tag{9}$$

对 $\boldsymbol{B}$ 进行基本行变换，使 $\boldsymbol{B}$ 中子块 $\boldsymbol{A}$ 变 $\boldsymbol{I}_n$, 即

$$\boldsymbol{B}\longrightarrow\widetilde{\boldsymbol{B}}=[\boldsymbol{I}_n\,|\,\widetilde{\boldsymbol{A}}]\tag{10}$$

这时第 2 个子块 $\widetilde{\boldsymbol{A}}$ 等于 $\boldsymbol{A}^{-1}$. 其原因是，只有利用矩阵 $\boldsymbol{A}^{-1}$ 才能实现 $\boldsymbol{A}\longrightarrow\boldsymbol{I}_n$ 的变换，即

$$\widetilde{\boldsymbol{B}}=\boldsymbol{A}^{-1}[\boldsymbol{A}\,|\,\boldsymbol{I}_n]=[\boldsymbol{I}_n\,|\,\boldsymbol{A}^{-1}]\tag{11}$$

(例)

试判别下列矩阵

$$\boldsymbol{A}=\begin{bmatrix}1&1\\1&2\end{bmatrix}\in R^{2\times 2}$$

是否是正则矩阵，若是，求其逆矩阵.

$$\boldsymbol{B}=\left[\begin{array}{cc|cc}1 & 1 & 1 & 0\\ 1 & 2 & 0 & 1\end{array}\right]\to\left[\begin{array}{cc|cc}1 & 1 & 1 & 0\\ 0 & 1 & -1 & 1\end{array}\right]\to\left[\begin{array}{cc|cc}1 & 0 & 2 & -1\\ 0 & 1 & -1 & 1\end{array}\right]$$

$$\Rightarrow \boldsymbol{A}^{-1}=\begin{bmatrix}2 & -1\\ -1 & 1\end{bmatrix}$$

因为 $\boldsymbol{B}$ 的 $\boldsymbol{A}$ 块能化成 $\boldsymbol{I}_2$, 所以 $\boldsymbol{A}$ 是正则矩阵.

B-I-3 矩阵的秩

矩阵的秩

利用 P-I-2 中讲过的方法可以由 $A\in F^{m\times n}$ 作出 1, 2, …, l 级(其中 $l=\min(m, n)$)子式.

[定义1] 当 $A\in F^{m\times n}$ 的 r 级子式中有不为 0 的子式存在(至少有一个)，而($r+1$)及($r+1$)级以上的子式全为 0 时,称 A 的行列式秩 (determinantal rank) 为 r, 记为

$$r=\operatorname{rank}A \tag{1}$$

也有很多表示成 $r=\rho(A)$.

由此定义,显然 $\operatorname{rank}A\leqslant\min(m, n)$. 而且,若所有 ($r+1$) 级子式全为 0, 由拉普拉斯展开式, 则所有($r+2$)级子式亦全为 0, 以下类推, 所有($r+1$)级及($r+1$)级以上的子式全为 0. 因此,研究矩阵的秩时只要确定所有的($r+1$)级子式全为 0 即可.

[例 1]

$$A=\begin{bmatrix}1&3&2&1\\2&0&1&2\\3&1&2&3\end{bmatrix}\in R^{3\times 4}$$

1 级子式: A 的各元素

2 级子式: 共有 $C_3^2\times C_4^2=18$ 个. 例如,取第 1, 2 行和第 1, 2 列

$$\det\begin{bmatrix}1&3\\2&0\end{bmatrix}=-6\neq 0$$

3 级子式: 共有 $C_3^3\times C_4^3=4$ 个.

$$\det\begin{array}{c}\begin{bmatrix}1&3&2\\2&0&1\\3&1&2\end{bmatrix}\\ \begin{array}{ccc}1&2&3\\ \text{列}&\text{列}&\text{列}\end{array}\end{array}=\det\begin{array}{c}\begin{bmatrix}1&3&1\\2&0&2\\3&1&3\end{bmatrix}\\ \begin{array}{ccc}1&2&4\\ \text{列}&\text{列}&\text{列}\end{array}\end{array}=\det\begin{array}{c}\begin{bmatrix}1&2&1\\2&1&2\\3&2&3\end{bmatrix}\\ \begin{array}{ccc}1&3&4\\ \text{列}&\text{列}&\text{列}\end{array}\end{array}=\det\begin{array}{c}\begin{bmatrix}3&2&1\\0&1&2\\1&2&3\end{bmatrix}\\ \begin{array}{ccc}2&3&4\\ \text{列}&\text{列}&\text{列}\end{array}\end{array}=0$$

因此, $\operatorname{rank}A=2$.

[例 2] 上述秩的定义对于多项式环 $P(s, R)$上的矩阵也适用.

例如,对于 $P(s, R)$环上的 2×2 矩阵

$$A=\begin{bmatrix}1+2s+s^2&0\\0&1-s\end{bmatrix} \tag{2}$$

由 $\det A=(1-s)(1+2s+s^2)$, 得 $\operatorname{rank}A=2$. 当给 s 以实数值时, 则变成 $A\in R^{2\times 2}$, 根据 s 取值不同,不一定是 $\operatorname{rank}A=2$. 事实上,当 $s=\pm 1$ 时, $\operatorname{rank}A=1$.

矩阵秩的性质(其 1)

由行列式的性质及矩阵秩的定义可直接得出下列性质.

在 $A\in F^{m\times n}$ 中,

(i) $\operatorname{rank}\boldsymbol{A}=\operatorname{rank}\boldsymbol{A}^T$

由此可见,关于行成立的性质,关于列亦成立.

(ii) $\boldsymbol{A}$ 的行(列)乘上任意的 $\alpha\in F$(而 $\alpha\neq 0$), $\operatorname{rank}\boldsymbol{A}$ 不变.

(iii) 将 $\boldsymbol{A}$ 的任意两行(列)互换, $\operatorname{rank}\boldsymbol{A}$ 不变.

(iv) 将 $\boldsymbol{A}$ 的某行(列)乘上 α 倍($\alpha\in F$)后加到另外一行(列), $\operatorname{rank}\boldsymbol{A}$ 也不变.

在上述性质中,显然(i)~(iii)不必解释. 现只对(iv)加以说明. 而且只要讨论了关于列的情况即足够了. 为了简单起见,现在我们只考虑将 $\boldsymbol{A}\in F^{m\times n}$ 的第2列乘上 α 加到第1列后得到的矩阵 $\widetilde{\boldsymbol{A}}$. 对于一般情况也可以进行完全同样的讨论.

设 $\boldsymbol{A}=[a_{ij}]$, 则

$$\widetilde{\boldsymbol{A}}=\begin{bmatrix} a_{11}+\alpha a_{12} & a_{12} & \cdots & a_{1n}\\ \vdots & \vdots & & \vdots\\ a_{m1}+\alpha a_{m2} & a_{m2} & \cdots & a_{mn}\end{bmatrix} \tag{3}$$

将 $\boldsymbol{A}$ 的第1列换成第2列,得到矩阵 $\widetilde{\widetilde{\boldsymbol{A}}}$,

$$\widetilde{\widetilde{\boldsymbol{A}}}\triangleq\begin{bmatrix} a_{12} & a_{12} & \cdots & a_{1n}\\ \vdots & \vdots & & \vdots\\ a_{m2} & a_{m2} & \cdots & a_{mn}\end{bmatrix} \tag{4}$$

设 $p\triangleq\operatorname{rank}\boldsymbol{A}+1$, 现来看 $\boldsymbol{A}$, $\widetilde{\boldsymbol{A}}$, $\widetilde{\widetilde{\boldsymbol{A}}}$ 的 p 级子式. 因为 $\boldsymbol{A}$ 和 $\widetilde{\boldsymbol{A}}$ 仅第1列不同, 故若 $j_1\neq 1$, 则

$$\det\widetilde{\boldsymbol{A}}\begin{pmatrix} i_1, & i_2, & \cdots, & i_p\\ j_1, & j_2, & \cdots, & j_p\end{pmatrix}=\det\boldsymbol{A}\begin{pmatrix} i_1, & i_2, & \cdots, & i_p\\ j_1, & j_2, & \cdots, & j_p\end{pmatrix}=0 \tag{5}$$

当 $j_1=1$ 时,根据 P-I-2 中讲过的行列式的性质(iv), 得

$$\begin{aligned}\det\widetilde{\boldsymbol{A}}\begin{pmatrix} i_1, & i_2, & \cdots, & i_p\\ 1, & j_2, & \cdots, & j_p\end{pmatrix}&=\det\boldsymbol{A}\begin{pmatrix} i_1, & i_2, & \cdots, & i_p\\ 1, & j_2, & \cdots, & j_p\end{pmatrix}+\alpha\det\widetilde{\widetilde{\boldsymbol{A}}}\begin{pmatrix} i_1, & i_2, & \cdots, & i_p\\ 1, & j_2, & \cdots, & j_p\end{pmatrix}\\ &=\alpha\det\widetilde{\widetilde{\boldsymbol{A}}}\begin{pmatrix} i_1, & i_2, & \cdots, & i_p\\ 1, & j_2, & \cdots, & j_p\end{pmatrix}\end{aligned} \tag{6}$$

若 $j_1=1$, $j_2=2$, 根据行列式的性质(vii), 得

$$\det\widetilde{\widetilde{\boldsymbol{A}}}\begin{pmatrix} i_1, & i_2, & i_3, & \cdots, & i_p\\ 1, & 2, & j_3, & \cdots, & j_p\end{pmatrix}=0 \tag{7}$$

在 $j_1=1$, $j_2\neq 2$ 的情况下

$$\det\widetilde{\widetilde{\boldsymbol{A}}}\begin{pmatrix} i_1, & i_2, & \cdots, & i_p\\ 1, & j_2, & \cdots, & j_p\end{pmatrix}=\det\boldsymbol{A}\begin{pmatrix} i_1, & i_2, & \cdots, & i_p\\ 2, & j_2, & \cdots, & j_p\end{pmatrix}=0 \tag{8}$$

由(7), (8)式可见, (6)式右边永远为0, 即

$$\begin{gathered}\det\widetilde{\boldsymbol{A}}\begin{pmatrix} i_1, & i_2, & \cdots, & i_p\\ j_1, & j_2, & \cdots, & j_p\end{pmatrix}=0\\ 1\leqslant i_1\leqslant i_2<\cdots<i_p\leqslant m\\ 1\leqslant j_1<j_2<\cdots<j_p\leqslant n\end{gathered} \tag{9}$$

因此, $\operatorname{rank}\widetilde{\boldsymbol{A}}\leqslant\operatorname{rank}\boldsymbol{A}$.

将 $\widetilde{\boldsymbol{A}}$ 第2列乘上 $(-\alpha)$ 加到第1列, 得到 $\boldsymbol{A}$. 根据上面的讨论, 知 $\operatorname{rank}\boldsymbol{A}\leqslant\operatorname{rank}\widetilde{\boldsymbol{A}}$. 因此, $\operatorname{rank}\widetilde{\boldsymbol{A}}=\operatorname{rank}\boldsymbol{A}$.

矩阵的秩和行(列)秩

在上一章中，将行(或列)标准形中不全为0的行(列)数定义为行(列)秩. 事实上与矩阵的秩一致，对此有下列重要定理.

[定理 1] 对于任意矩阵 $\boldsymbol{A}\in F^{m\times n}$, 行秩、列秩和行列式秩均完全一致.

(证明) 设 $\boldsymbol{A}$ 的标准形为 $\widetilde{\boldsymbol{A}}_r$, 根据矩阵秩的性质 (ii)～(iv), 进行基本行变换并不改变矩阵的秩 r_1. 当然，进行基本行变换后行秩 r_2 亦不会变化. 比较 r_1 和 r_2 对 $\widetilde{\boldsymbol{A}}_r$ 进行即可. 由行标准形的性质可知，在 $\widetilde{\boldsymbol{A}}_r$ 上作 $(r_2+1)\times(r_2+1)$ 阶子矩阵，必然包含元素全为0的行. 由 P-I-1 中行列式的性质 (ii) 可见，$\widetilde{\boldsymbol{A}}_r$ 的 (r_2+1) 级子式全为0, 所以 $r_1\leqslant r_2$. 在 $\widetilde{\boldsymbol{A}}_r$ 中，除了第1～第 r_2 行外，消去其它各行; 除了第 k_1, … 第 kr_2 列外，消去其它各列，则得到的 $r_2\times r_2$ 阶子矩阵是 $\boldsymbol{I}_{r_2}$. 显然，其行列式不为0, 则 $r_1\geqslant r_2$. 因此，$r_1=r_2$. 对于列秩也是一样.

根据定理1, 以前所定义的秩都完全一致，显然没有区别的必要. 因此今后不加以区分，只称秩.

秩的性质(其 2)[22]

将 B-I-2 定理的[系 3]和基本行 (列) 变换后矩阵的秩不变的性质结合起来，便得到下列应用很方便的结果.

[系 1]

设 $\boldsymbol{P}$ 为任意 $m\times m$ 正则矩阵，$\boldsymbol{Q}$ 为任意 $n\times n$ 正则矩阵，$\boldsymbol{A}$ 为任意 $m\times n$ 矩阵(都是体 F 上的矩阵)，则

$$\operatorname{rank}\boldsymbol{A}=\operatorname{rank}\boldsymbol{PA}=\operatorname{rank}\boldsymbol{AQ}=\operatorname{rank}\boldsymbol{PAQ}$$

[定义 2] 在[系 1]的意义上，设 $\boldsymbol{A}\in F^{m\times n}$, $\boldsymbol{B}\in F^{m\times n}$ ，有正则矩阵，$\boldsymbol{P}\in F^{m\times m}$, $\boldsymbol{Q}\in F^{n\times n}$ 存在，使得

$$\boldsymbol{B}=\boldsymbol{PAQ}\tag{10}$$

则称 $\boldsymbol{A}$, $\boldsymbol{B}$互相等价(equivalent).

特别是，当 $m=n$, $\boldsymbol{Q}=\boldsymbol{P}^{-1}$ 时，即对于某个正则矩阵 $\boldsymbol{P}\in F^{n\times n}$, $\boldsymbol{A}$, $\boldsymbol{B}\in F^{n\times n}$ 之间存在

$$\boldsymbol{B}=\boldsymbol{PAP}^{-1}\tag{11}$$

时，称为 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 相似(similar).

(问题 1) 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 相似，试证明

$$\det\boldsymbol{A}=\det\boldsymbol{B}\tag{12}$$

$$\operatorname{tr}\boldsymbol{A}=\operatorname{tr}\boldsymbol{B}\tag{13}$$

(问题 2) 对于任意矩阵 $\boldsymbol{A}\in F^{m\times n}$, 试证明

$$\operatorname{rank}\boldsymbol{A}=\operatorname{rank}\boldsymbol{AA}^T=\operatorname{rank}\boldsymbol{A}^T\boldsymbol{A}\tag{14}$$

(提示: 设与 $\boldsymbol{A}$ 行等价的行标准形为 $\widetilde{\boldsymbol{A}}_r=\boldsymbol{PA}$, 此时，若 $\boldsymbol{A}^T$ 的列等价标准形为 $\widetilde{\boldsymbol{A}}_c$, 则 $\widetilde{\boldsymbol{A}}_c=\widetilde{\boldsymbol{A}}_r^T=\boldsymbol{A}^T\boldsymbol{P}^T$. 因此，$\operatorname{rank}\boldsymbol{A}=\operatorname{rank}\widetilde{\boldsymbol{A}}_r\widetilde{\boldsymbol{A}}_c=\operatorname{rank}\boldsymbol{AA}^T$).

对于这种特殊情况，$\operatorname{rank}\boldsymbol{A}=n\Leftrightarrow\det\boldsymbol{A}^T\boldsymbol{A}\neq0$, 意味着 $\operatorname{rank}\boldsymbol{A}=m\Leftrightarrow\det\boldsymbol{AA}^T\neq0$.

(问题 3) 对于任意矩阵 $\boldsymbol{A}\in F^{m\times n}$, $\boldsymbol{B}\in F^{n\times m}$, 试证明

$$\mathrm{rank}\boldsymbol{AB}=\mathrm{rank}\boldsymbol{BA} \tag{15}$$

是否成立．若不成立，试举出相反的例子．

根据定理 1，矩阵的秩等于其行秩、列秩．对于分块矩阵显然有如下关系．将 $\boldsymbol{A}\in F^{m\times n}$ 无论分成 $\boldsymbol{A}=[\boldsymbol{A}_1|\boldsymbol{A}_2]$ 或 $\boldsymbol{A}=\left[\frac{\boldsymbol{A}_1}{\boldsymbol{A}_2}\right]$，下式均成立

$$\max\{\mathrm{rank}\boldsymbol{A}_1,\ \mathrm{rank}\boldsymbol{A}_2\}\leqslant\mathrm{rank}\boldsymbol{A}\leqslant\min\{m,\ n,\ \mathrm{rank}\boldsymbol{A}_1+\mathrm{rank}\boldsymbol{A}_2\} \tag{16}$$

如令式中 $\boldsymbol{A}_2=\boldsymbol{0}$（零矩阵），因为零矩阵的秩显然是零，则

$$\mathrm{rank}[\boldsymbol{A}|\boldsymbol{0}]=\mathrm{rank}\left[\frac{\boldsymbol{A}}{\boldsymbol{0}}\right]=\mathrm{rank}\boldsymbol{A}. \tag{17}$$

（问题 4） 设 $\boldsymbol{A}\in F^{m\times n}$，$\boldsymbol{B}\in F^{m\times n}$，试证明

$$\mathrm{rank}(\boldsymbol{A}+\boldsymbol{B})\leqslant\mathrm{rank}(\boldsymbol{A})+\mathrm{rank}(\boldsymbol{B}) \tag{18}$$

及下列关系是否始终成立

$$\mathrm{rank}(\boldsymbol{A}+\boldsymbol{A})=\mathrm{rank}\boldsymbol{A} \tag{19}$$

（问题 5） 对于下列分块三角形矩阵

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1n}\\ \boldsymbol{0} & \boldsymbol{A}_{22} & \cdots & \boldsymbol{A}_{2n}\\ \vdots & \vdots & & \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{A}_{nn}\end{bmatrix}, \tag{20}$$

试证明 $\mathrm{rank}\boldsymbol{A}\geqslant\sum\limits_{i=1}^{n}\mathrm{rank}\boldsymbol{A}_{ii}$，特别是对于分块对角矩阵

$$\boldsymbol{A}=\boldsymbol{A}_{11}\dotplus\boldsymbol{A}_{22}\dotplus\cdots\dotplus\boldsymbol{A}_{nn} \tag{21}$$

试证明

$$\mathrm{rank}\boldsymbol{A}=\sum_{i=1}^{n}\mathrm{rank}\boldsymbol{A}_{ii} \tag{22}$$

在系统理论中，如何确定矩阵 $\boldsymbol{A}\in F^{m\times n}$，$\boldsymbol{B}\in F^{n\times q}$ 之积的秩常是个问题．我们知道，对此有下列西勒维斯特不等式（Sylvester's inequality）[23]

$$\mathrm{rank}\boldsymbol{A}+\mathrm{rank}\boldsymbol{B}-n\leqslant\mathrm{rank}(\boldsymbol{AB})\leqslant\min(\mathrm{rank}\boldsymbol{A},\ \mathrm{rank}\boldsymbol{B}) \tag{23}$$

在第 II 部分中，这个结论可以从几何学的观点得出，这里只是把结果提一下．

行（列）向量的线性独立和矩阵的秩

如前所述，体 F 上的 $m\times n$ 矩阵可以用分割成行或列的形式表示

$$\boldsymbol{A}=\begin{bmatrix}a_{11} & \cdots & a_{1n}\\ \vdots & & \vdots\\ a_{m1} & \cdots & a_{mn}\end{bmatrix}=[\boldsymbol{a}_1,\ \boldsymbol{a}_2,\ \cdots,\ \boldsymbol{a}_n]=\begin{bmatrix}\tilde{\boldsymbol{a}}_1\\ \tilde{\boldsymbol{a}}_2\\ \vdots\\ \tilde{\boldsymbol{a}}_m\end{bmatrix} \tag{24}$$

式中 $\boldsymbol{a}_j(j=1,\ 2,\ \cdots,\ n)$ 是 $\boldsymbol{A}$ 的第 j 列，$\tilde{\boldsymbol{a}}_i(i=1,\ 2,\ \cdots,\ m)$ 是 $\boldsymbol{A}$ 的第 i 行，即

$$\boldsymbol{a}_j\triangleq\begin{bmatrix}a_{1j}\\ a_{2j}\\ \vdots\\ a_{mj}\end{bmatrix},\ \tilde{\boldsymbol{a}}_i\triangleq[a_{i1},\ a_{i2},\ \cdots,a_{in}] \tag{25}$$

由此可见，$\boldsymbol{A}$ 的列可以看成向量空间 F^m 中的向量；$\boldsymbol{A}$ 的行可以看成向量空间 F^n 中的向量．

这样看时，$\boldsymbol{A}$ 的行(列)向量和前节所定义的 $\boldsymbol{A}$ 的秩之间有什么关系呢？若将秩看成由 $\boldsymbol{A}$ 决定的数值，则其与 $\boldsymbol{A}$ 的整个行(列)向量组的性质（即不是某一个向量的性质）相对应．下面的定理回答了这个问题．

[定理 2]

设 $\boldsymbol{A}\in F^{m\times n}$ 的第 i 行为 $\tilde{\boldsymbol{a}}_i$，第 j 列为 $\boldsymbol{a}_j$，即

$$\boldsymbol{A}=[a_{ij}]=[\boldsymbol{a}_1,\ \boldsymbol{a}_2,\ \cdots,\ \boldsymbol{a}_n]=\begin{bmatrix}\tilde{\boldsymbol{a}}_1\\ \tilde{\boldsymbol{a}}_2\\ \vdots\\ \tilde{\boldsymbol{a}}_m\end{bmatrix} \tag{26}$$

式中 $\boldsymbol{a}_j\in F^m$，$\tilde{\boldsymbol{a}}_i\in F^n$．

所谓 $\operatorname{rank}\boldsymbol{A}=r$，和下列两件事情等价：$m$ 个向量 $\{\tilde{\boldsymbol{a}}_1,\ \tilde{\boldsymbol{a}}_2,\ \cdots,\ \tilde{\boldsymbol{a}}_m\}$ 中有 r 个线性独立的向量，其它 $(m-r)$ 个向量可以用其线性组合表示；此外，n 个向量 $\{\boldsymbol{a}_1,\ \boldsymbol{a}_2,\ \cdots,\ \boldsymbol{a}_n\}$ 中有 r 个线性独立的向量，其它 $(n-r)$ 个向量可以用其线性组合表示．

在证明该定理之前，先来证明下列引理．

[引理]　设 $\boldsymbol{v}_i\in F^n$，$i=1,\ 2,\ \cdots,\ r$，则对于任意正则矩阵 $\boldsymbol{P}\in F^{n\times n}$

$\boldsymbol{v}_i(i=1,\ 2,\ \cdots,\ r)$ 线性独立 $\Leftrightarrow \boldsymbol{P}\boldsymbol{v}_i(i=1,\ 2,\ \cdots,\ r)$ 线性独立

（证明）　先假定 $\boldsymbol{v}_i$ 线性独立，根据线性独立的定义，若令 $\alpha_i\in F(i=1,\ 2,\ \cdots,\ r)$，则

$$\alpha_1\boldsymbol{v}_1+\cdots+\alpha_r\boldsymbol{v}_r=\boldsymbol{0}\Leftrightarrow\alpha_1=\cdots=\alpha_r=0,$$

因 $\boldsymbol{P}$ 为正则矩阵，故

$$\boldsymbol{P}(\alpha_1\boldsymbol{v}_1+\cdots+\alpha_r\boldsymbol{v}_r)=\boldsymbol{0}\Leftrightarrow\alpha_1\boldsymbol{v}_1+\cdots+\alpha_r\boldsymbol{v}_r=\boldsymbol{0}$$

容易看到，纯量和矩阵 $\boldsymbol{P}$ 相乘，其顺序可以交换，因此

$$\boldsymbol{P}(\alpha_1\boldsymbol{v}_1+\cdots+\alpha_r\boldsymbol{v}_r)=\alpha_1\boldsymbol{P}\boldsymbol{v}_1+\cdots+\alpha_r\boldsymbol{P}\boldsymbol{v}_r$$

由以上可见

$$\alpha_1\boldsymbol{P}\boldsymbol{v}_1+\cdots+\alpha_r\boldsymbol{P}\boldsymbol{v}_r=\boldsymbol{0}\Leftrightarrow\alpha_1=\cdots=\alpha_r=0$$

于是证明了 $\boldsymbol{P}\boldsymbol{v}_i$ 线性独立．反之亦然．

（定理 2 的证明）

因 $\operatorname{rank}\boldsymbol{A}=\operatorname{rank}\boldsymbol{A}^T$，故只对于列证明即可．首先证明“$\operatorname{rank}\boldsymbol{A}=r\Rightarrow\boldsymbol{A}$ 的列中有 r 个线性独立的向量；其余向量均可用其线性组合表示”．根据引理，因为基本变换矩阵都是正则矩阵，故对矩阵 $\boldsymbol{A}$ 进行基本变换毫不影响 $\boldsymbol{A}$ 的列的线性独立．因此，也可以代替 $\boldsymbol{A}$ 考查 $\boldsymbol{A}$ 的标准形 $\tilde{\boldsymbol{A}}_r$．

显然，$\tilde{\boldsymbol{A}}_r$ 的第 k_1 列，第 k_2 列，…，第 k_r 列是线性独立的，所以在 $\tilde{\boldsymbol{A}}_r$ 的列中有 r 个线性独立的向量，而 $\tilde{\boldsymbol{A}}_r$ 的其余 $(n-r)$ 个列，即从下边数 $(n-r)$ 个元素都是 0，即

$$\boldsymbol{a}_j=\begin{bmatrix}c_{j1}\\ c_{j2}\\ \vdots\\ c_{jr}\\ 0\\ \vdots\\ 0\end{bmatrix}\begin{matrix}\left.\vphantom{\begin{matrix}c\\c\\ \vdots\\c\end{matrix}}\right\}r\\ \left.\vphantom{\begin{matrix}0\\ \vdots\\0\end{matrix}}\right\}n-r\end{matrix} \tag{27}$$

因此，可用下列线性组合表示

$$\boldsymbol{a}_j = c_{j1}\boldsymbol{a}_{k1} + c_{j2}\boldsymbol{a}_{k2} + \cdots + c_{jr}\boldsymbol{a}_{kr} \tag{28}$$

其次，我们来证明“线性独立的列向量有 r 个，其他各个列向量可用其线性组合表示 $\Rightarrow \mathrm{rank}\boldsymbol{A} = r$”.

设 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_r$ 是线性独立的，及

$$\tilde{\boldsymbol{A}} \triangleq [\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_r] \in F^{m\times r} \tag{29}$$

根据假定，$\boldsymbol{a}_{r+1}, \boldsymbol{a}_{r+2}, \cdots, \boldsymbol{a}_n$ 可以用上述列向量的线性组合表示．例如，$\boldsymbol{a}_{r+1}$ 可以表示成

$$\boldsymbol{a}_{r+1} = \alpha_{11}\boldsymbol{a}_1 + \alpha_{21}\boldsymbol{a}_2 + \cdots + \alpha_{r1}\boldsymbol{a}_r \tag{30}$$

由此可见，反复使用秩的性质(iv)，即

$$\mathrm{rank}[\tilde{\boldsymbol{A}} \mid \boldsymbol{a}_{r+1}] = \mathrm{rank}[\tilde{\boldsymbol{A}} \mid \alpha_{11}\boldsymbol{a}_1 + \cdots + \alpha_{r1}\boldsymbol{a}_r] = \mathrm{rank}[\tilde{\boldsymbol{A}} \mid \boldsymbol{0}] = \mathrm{rank}\tilde{\boldsymbol{A}} \tag{31}$$

对于其它列也是同样，最后可得

$$\mathrm{rank}\boldsymbol{A} = \mathrm{rank}[\tilde{\boldsymbol{A}} \mid \boldsymbol{a}_{r+1} \mid \boldsymbol{a}_{r+2} \mid \cdots \mid \boldsymbol{a}_n] = \mathrm{rank}[\tilde{\boldsymbol{A}} \mid \boldsymbol{0} \mid \boldsymbol{0} \mid \cdots \mid \boldsymbol{0}] = \mathrm{rank}\tilde{\boldsymbol{A}} \tag{32}$$

所以 $\mathrm{rank}\boldsymbol{A} = \mathrm{rank}\tilde{\boldsymbol{A}} \leqslant r$．若 $\mathrm{rank}\boldsymbol{A} < r$，则 $\boldsymbol{A}$ 的列中所包含的线性独立向量的最大数小于 r(使用证明的前半部分)，与假定相反．因此，$\mathrm{rank}\boldsymbol{A} = r$.

[例 3]

$$\boldsymbol{A} = \begin{bmatrix} 1 & 3 & 2 & 1 \\ 2 & 0 & 1 & 2 \\ 3 & 1 & 2 & 3 \end{bmatrix} \in R^{3\times 4}$$

在例 1 中已经证明，$\mathrm{rank}\boldsymbol{A} = 2$．因位于 $\boldsymbol{A}$ 中左上角的 2 级子式

$$\det\begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix} = -6 \neq 0$$

则第 1 列和第 2 列 (或第 1 行和第 2 行) 是线性独立的．第 3 列，第 4 列可以用第 1 列、第 2 列线性表出

$$\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + 0\begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$$

[例 4]

$$\boldsymbol{A} = \begin{bmatrix} \dfrac{1}{s+1} & \dfrac{s+2}{(s+1)(s+5)} \\ \dfrac{1}{s+2} & \dfrac{1}{s+5} \end{bmatrix}$$

$\boldsymbol{A}$ 是实系数有理函数体上的 2×2 阶方阵，即 $F = Q(s, R)$．$\det\boldsymbol{A} = 0$，这意味着第 1 列和第 2 列(第 1 行和第 2 行)在体 $Q(s, R)$ 上的向量空间 $Q(s, R)\times Q(s, R)$ (即 F^2) 中线性相关．如在 B-0-2 中所述，显然能够用 $Q(s, R)\times Q(s, R)$ 构成体 R 上的向量空间．在这种情况下，因数所属的体(R)和向量部分所属的体 ($Q(s, R)$)不同，所以矩阵的秩和矩阵的行(列)线性独立无关．当看作是体 R 上的向量空间 $Q(s, R)\times Q(s, R)$ 时，第 1 列和第 2 列是线性独立的.

[例 5]

$$A=\begin{bmatrix}\dfrac{1}{s+1} & \dfrac{2}{s+1}\\ 1 & 2\end{bmatrix},\ F=Q(s,\ R)$$

$\det A=0$, 对于 F^2 向量空间，第 1, 2 行(列)是线性相关的．当看成是 R 上的向量空间 $Q(s,\ R)\times Q(s,\ R)$时，显然第 1 列和第 2 列也是线性相关的(2×第 1 列-第 2 列$=\mathbf{0}$)．按后一看法，第 1 行和第 2 行是线性独立的．

(注意)

如在上面诸例中所见，当把矩阵 $\boldsymbol{A}\in F^{m\times n}$ 的行(列)看作向量空间 $F^m(F^n)$ 中的向量时，可以利用 $\boldsymbol{A}$ 的秩判别这些向量的线性独立．另外，向量空间不仅限于 $F^m(F^n)$，可以更普遍地定义出各种不同的向量空间．对于这些向量空间中的向量，B-0-2 中讲过的线性独立的定义仍然适用．不过要注意，不是在任何情况下都可以由矩阵的秩判别是否线性独立．

秩为 r 的矩阵的性质

设矩阵 $\boldsymbol{A}\in F^{m\times n}$ 的秩为 r，现来研究一下 $\boldsymbol{A}$ 的性质(不变成特殊形的分解等)．其结果与下一章中线性方程组的解法及以后将要讲的一般化逆矩阵的求法有关，而且也是很重要的．

[定理 3]

设矩阵 $A\in F^{m\times n}$ 的秩为 r，则有满足

$$\widetilde{\boldsymbol{A}}=\boldsymbol{P}_m\boldsymbol{A}\boldsymbol{P}_n,\ \widetilde{\boldsymbol{A}}=\begin{array}{c}\begin{array}{cc}\ r & \ n-r\end{array}\\ \left[\begin{array}{c|c}\overbrace{\widetilde{\boldsymbol{A}}_{11}} & \overbrace{\widetilde{\boldsymbol{A}}_{12}}\\ \hline \widetilde{\boldsymbol{A}}_{21} & \widetilde{\boldsymbol{A}}_{22}\end{array}\right]\end{array}\begin{array}{l}\\ )r\\ )m-r\end{array}$$

$$\det\widetilde{\boldsymbol{A}}_{11}\neq 0,\ \widetilde{\boldsymbol{A}}_{22}=\widetilde{\boldsymbol{A}}_{21}\widetilde{\boldsymbol{A}}_{11}^{-1}\widetilde{\boldsymbol{A}}_{12} \tag{33}$$

的顺序矩阵 $\boldsymbol{P}_m$, $\boldsymbol{P}_n$ 存在．

(证明) 若 $\boldsymbol{A}$ 的秩为 r，则其 r 阶子式中至少有一个不为 0 的子式存在．交换 $\boldsymbol{A}$ 的行列，使与不为 0 的 r 级子式相对应的 $r\times r$ 阶矩阵移到左上角，结果在所得到的矩阵 $\widetilde{\boldsymbol{A}}$ 中，有 $\det\widetilde{\boldsymbol{A}}_{11}\neq 0$．因 $\widetilde{\boldsymbol{A}}$ 的第 1 行～第 r 行是线性独立的，根据定理 2，则第 $r+1$ 行～第 m 行可以用第 1 行～第 r 行的线性组合表示．换句话说，有同时满足下列两式

$$\boldsymbol{X}\widetilde{\boldsymbol{A}}_{11}=\widetilde{\boldsymbol{A}}_{21},\ \boldsymbol{X}\widetilde{\boldsymbol{A}}_{12}=\widetilde{\boldsymbol{A}}_{22} \tag{34}$$

的 $(m-r)\times r$ 矩阵 $\boldsymbol{X}$ 存在．因 $\boldsymbol{X}=\widetilde{\boldsymbol{A}}_{21}\widetilde{\boldsymbol{A}}_{11}^{-1}$，很容易看到，满足其中第一式的矩阵 $\boldsymbol{X}$ 仅只有一个．因此由第二式得

$$\widetilde{\boldsymbol{A}}_{22}=\widetilde{\boldsymbol{A}}_{21}\widetilde{\boldsymbol{A}}_{11}^{-1}\widetilde{\boldsymbol{A}}_{12} \tag{35}$$

同样，因 $\widetilde{\boldsymbol{A}}$ 的第 1 列～第 r 列是线性独立的，则有同时满足下列两式

$$\widetilde{\boldsymbol{A}}_{11}\boldsymbol{Y}=\widetilde{\boldsymbol{A}}_{12},\ \widetilde{\boldsymbol{A}}_{21}\boldsymbol{Y}=\widetilde{\boldsymbol{A}}_{22} \tag{36}$$

的 $r\times(n-r)$ 矩阵 $\boldsymbol{Y}$ 存在．满足其中第一式的 $\boldsymbol{Y}$ 仅只有一个，$\boldsymbol{Y}=\widetilde{\boldsymbol{A}}_{11}^{-1}\widetilde{\boldsymbol{A}}_{12}$．将其代入第二式，则又得到(35)式．

如 B-I-3 中所述，矩阵 $\boldsymbol{A}\in F^{m\times n}$ 的秩 r 不能超过 $\min(m,\ n)$．当 $r=\min(m,\ n)$时，矩阵 $\boldsymbol{A}$ 具有最大秩，或称为最大秩的矩阵．把矩阵 $\boldsymbol{A}$ 的行看成 F^n 的向量，列看成 F^m 的

向量时，由定理2可见，在最大秩矩阵 $\boldsymbol{A}$ 中，(i) 若 $m \leqslant n$，则所有的行都是线性独立的，(ii) 若 $m \geqslant n$，则所有的列都是线性独立的．下列性质是很重要的．

［定理4］ 设矩阵 $\boldsymbol{A} \in F^{m \times n}$ 的秩为 r，则可将 $\boldsymbol{A}$ 分解成均具有最大秩的矩阵 $\boldsymbol{B} \in F^{m \times r}$，$\boldsymbol{C} \in F^{r \times n}$ 之积，即．

$$\boldsymbol{A}=\boldsymbol{BC},\ \operatorname{rank}\boldsymbol{A}=\operatorname{rank}\boldsymbol{B}=\operatorname{rank}\boldsymbol{C}=r \tag{37}$$

（证明） 如在定理2的证明中所述，对于基本变换，矩阵列的线性独立、线性相关的性质不变．关于和 $\boldsymbol{A}$ 行等价的矩阵 $\tilde{\boldsymbol{A}}$，下列两式

(a) $\boldsymbol{A}$ 的第 i_1 列，第 i_2 列，…，第 i_q 列线性独立 $\Leftrightarrow \tilde{\boldsymbol{A}}$ 的第 i_1 列，第 i_2 列，…，第 i_q 列线性独立

(b) $\boldsymbol{A}$ 的第 j 列可以用 $\boldsymbol{A}$ 的第 i_1 列，第 i_2 列，…，第 i_q 列的线性组合表示

$$\boldsymbol{a}_j=c_{j1}\boldsymbol{a}_{i1}+c_{j2}\boldsymbol{a}_{i2}+\cdots+c_{jq}\boldsymbol{a}_{iq}$$

$\Leftrightarrow \tilde{\boldsymbol{A}}$ 的第 j 列可以用 $\tilde{\boldsymbol{A}}$ 的 i_1 列，第 i_2 列，…，第 i_q 列的线性组合表示

$$\tilde{\boldsymbol{a}}_j=c_{j1}\tilde{\boldsymbol{a}}_{i1}+c_{j2}\tilde{\boldsymbol{a}}_{i2}+\cdots+c_{jq}\tilde{\boldsymbol{a}}_{iq} \tag{38}$$

对于任意的 $1 \leqslant i_1,\ i_2,\ \cdots,\ i_q,\ q,\ j \leqslant n$ 均成立．

对于 $\tilde{\boldsymbol{A}}$，现来看 $\boldsymbol{A}$ 的标准形 $\tilde{\boldsymbol{A}}_r$．因为显然 $\tilde{\boldsymbol{A}}_r$ 的第 k_1 列，…，第 k_r 列是线性独立的，则由上面的性质(a)可知，$\boldsymbol{A}$ 的第 k_1 列，…，第 k_r 列也是线性独立的．因此，除了 $\boldsymbol{A}$ 第 k_1 列，…，第 k_r 列外，若其它各列都消去，设所得到的矩阵为 $\boldsymbol{B}$，则 $\boldsymbol{B}$ 是具有最大秩的 $m \times r$ 矩阵．

$\tilde{\boldsymbol{A}}_r$ 的第 j 列可以表示成

$$\tilde{\boldsymbol{a}}_j=\begin{bmatrix} c_{j1} \\ \vdots \\ c_{jp} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix} c \\ \vdots \\ c \end{matrix}}\right\} p \\ \left.\vphantom{\begin{matrix} 0 \\ \vdots \\ 0 \end{matrix}}\right\} n-p \end{matrix} \tag{39}$$

式中 $k_p<j<k_{p+1}\ (1 \leqslant p \leqslant r)$，因此它可以用下列线性组合表示

$$\tilde{\boldsymbol{a}}_j=c_{j1}\tilde{\boldsymbol{a}}_{k1}+\cdots+c_{jp}\tilde{\boldsymbol{a}}_{kp} \tag{40}$$

由此及性质 (b)，$\boldsymbol{A}$ 的第 j 列也可以用同样的线性组合表示．其次，设除了位于 $\tilde{\boldsymbol{A}}_r$ 下侧 $(n-r)$ 个 $\boldsymbol{0}$ 行以外的 $r \times n$ 矩阵为 $\boldsymbol{C}$，显然 $\boldsymbol{C}$ 是具有最大秩的 $r \times n$ 矩阵．当 $k_p<j<k_{p+1}$ 时，则

$$\boldsymbol{C}\text{ 的第 } j \text{ 列}=\begin{bmatrix} c_{j1} \\ \vdots \\ c_{jp} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix} c \\ \vdots \\ c \end{matrix}}\right\} p \\ \left.\vphantom{\begin{matrix} 0 \\ \vdots \\ 0 \end{matrix}}\right\} r-p \end{matrix} \tag{41}$$

而且，因 $\boldsymbol{B}=[\boldsymbol{a}_{k1},\ \boldsymbol{a}_{k2},\ \cdots,\ \boldsymbol{a}_{kr}]$，则

$$[\boldsymbol{A}\text{ 的第 } j \text{ 列}]=\boldsymbol{B}\times[\boldsymbol{C}\text{ 的第 } j \text{ 列}]$$

式中 $j=k_1,\ k_2,\ \cdots,\ k_r$．对于 $j<k_1$，$j>k_r$ 也是同样．最后，得 $\boldsymbol{A}=\boldsymbol{BC}$．

（问题6） 在定理3中的矩阵 $\tilde{\boldsymbol{A}}$ 可以分解成

$$\tilde{A}=\left[\frac{\tilde{A}_{11}}{\tilde{A}_{12}}\right][I|\ \tilde{A}_{11}^{-1}\tilde{A}_{12}]=\left[\frac{I}{\tilde{A}_{21}\tilde{A}_{11}^{-1}}\right][\tilde{A}_{11}|\tilde{A}_{12}] \tag{42}$$

试用该式证明定理 4.

在定理 4 中将矩阵 $A\in F^{m\times n}$ 按 (37) 式分解. 一般, 对于和 (37) 式不同形式的 A 分解

$$A=\tilde{B}\tilde{C},\ \tilde{B}\in F^{m\times p},\ \tilde{C}\in F^{p\times n} \tag{43}$$

下列定理成立.

[定理 5] 秩为 r 的矩阵 A 可以按 (37) 式分解成具有最大秩的矩阵 B, C, 也可以按 (43) 式分解成和 B, C 不同的某个矩阵组 $\tilde{B}$, $\tilde{C}$. 这时, 一般必须 $p\geqslant r$, 而且有满足

$$\begin{aligned} C&=Q_1\tilde{C}\\ B&=\tilde{B}Q_2 \end{aligned} \tag{44}$$

的 $Q_1\in F^{r\times p}$, $Q_2\in F^{p\times r}$ 存在. 若 $p=r$, 则 $\tilde{B}$, $\tilde{C}$ 具有最大秩, Q_1, Q_2 是正则矩阵, 而且 $Q_1^{-1}=Q_2$.

(证明) 由 (23) 式可知

$$r=\mathrm{rank}A\leqslant\min(\mathrm{rank}\tilde{B},\ \mathrm{rank}\tilde{C})\leqslant\min(m,\ n,\ p) \tag{45}$$

则 $p\geqslant r$, 由 (37), (43) 式得

$$BC=\tilde{B}\tilde{C}. \tag{46}$$

将等式两边左乘以 B^T, 根据问题 2 中讲过的性质, 并考虑到 B^TB 是正则矩阵, 故

$$C=(B^TB)^{-1}B^T\tilde{B}\tilde{C} \tag{47}$$

于是 (44) 式的第一式得到证明. 当 $p=r$ 时, 因 $\mathrm{rank}A=r\leqslant\min(m,\ n)$, 则由 (45) 式得

$$r=\mathrm{rank}A\leqslant\min(\mathrm{rank}\tilde{B},\ \mathrm{rank}\tilde{C})\leqslant\min(m,\ n,\ r)=r \tag{48}$$

且因 $\mathrm{rank}\tilde{B}\leqslant\min(m,\ r)=r$, $\mathrm{rank}\tilde{C}\leqslant\min(n,\ r)=r$, 则说明 $\mathrm{rank}\tilde{B}=\mathrm{rank}\tilde{C}=r$, 即 $\tilde{B}$, $\tilde{C}$ 都具有最大秩. 因此, 和由 (46) 式推导 (47) 式的过程一样, 得到

$$\tilde{C}=(\tilde{B}^T\tilde{B})^{-1}\tilde{B}^TBC \tag{49}$$

$$\therefore\quad C=[(B^TB)^{-1}B^T\tilde{B}][(\tilde{B}^T\tilde{B})^{-1}\tilde{B}^TB]C$$

将等式两边右乘以 C^T, 由问题 2 中所讲过的性质得知, CC^T 是正则矩阵, 则

$$I=[(B^TB)^{-1}B^T\tilde{B}][(\tilde{B}^T\tilde{B})^{-1}\tilde{B}^TB]$$

$$\therefore\quad [(B^TB)^{-1}B^T\tilde{B}]^{-1}=[(\tilde{B}^T\tilde{B})^{-1}\tilde{B}^TB] \tag{50}$$

说明 (44) 式中的 Q_1 是正则矩阵. 对于 (44) 式的第二式也可按同样办法证明. 在 $p=r$ 的情况下, 仍然根据问题 2, $\tilde{B}^T\tilde{B}$, $\tilde{C}\tilde{C}^T$ 也是正则矩阵, 由 (43), (44) 式得

$$\tilde{B}\tilde{C}=A=BC=\tilde{B}Q_2Q_1\tilde{C}$$

注意等式的左边和右边, 两边都左乘以 $\tilde{B}^T$, 右乘以 $\tilde{C}^T$, 得

$$(\tilde{B}^T\tilde{B})(\tilde{C}\tilde{C}^T)=(\tilde{B}^T\tilde{B})Q_2Q_1(\tilde{C}\tilde{C}^T)$$

$$\therefore\quad I=Q_2Q_1 \tag{51}$$

$$\therefore\quad Q_1^{-1}=Q_2 \tag{52}$$

[系 1] 当 A 分解成下列因式

$$A=BC=\tilde{B}\tilde{C}$$

时, 若 B, C, $\tilde{B}$, $\tilde{C}$ 都具有最大秩, 则

$$C^T(CC^T)^{-1}(B^TB)^{-1}B^T=\tilde{C}^T(\tilde{C}\tilde{C}^T)^{-1}(\tilde{B}^T\tilde{B})^{-1}\tilde{B}^T.$$

B-I-4 线性方程组

本章将介绍线性方程组的基本内容. 设 $x_1, x_2, \cdots, x_n$ 为未知数，方程式的系数是体 F 上的元. 首先讨论未知数的个数和方程式的个数相等的情况，即

$$\begin{aligned} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n&=b_1\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n&=b_2\\ &\vdots\\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n&=b_n \end{aligned} \tag{1a}$$

若利用矩阵符号

$$\boldsymbol{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{bmatrix},\quad \boldsymbol{b}=\begin{bmatrix} b_1\\ b_2\\ \vdots\\ b_n\end{bmatrix},\quad \boldsymbol{x}=\begin{bmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{bmatrix}$$

则上述方程式可表示成

$$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b} \tag{1b}$$

当 $\det\boldsymbol{A}\neq 0$ 时，(1b)式(即(1a)式)的解 $\boldsymbol{x}$ 是唯一存在的，由下式给出

$$\boldsymbol{x}=\boldsymbol{A}^{-1}\boldsymbol{b} \tag{2}$$

因 $\boldsymbol{A}$ 的逆矩阵 $\boldsymbol{A}^{-1}$ 可以用其代数余子式表示

$$\boldsymbol{A}^{-1}=\frac{1}{\det\boldsymbol{A}}\begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1}\\ A_{12} & A_{22} & \cdots & A_{n2}\\ \vdots & \vdots & & \vdots\\ A_{1n} & A_{2n} & \cdots & A_{nn}\end{bmatrix}$$

则(2)式可以写成

$$\begin{bmatrix} x_1\\ \vdots\\ x_n\end{bmatrix}=\frac{1}{\det\boldsymbol{A}}\begin{bmatrix} A_{11}b_1+A_{21}b_2+\cdots+A_{n1}b_n\\ \vdots\\ A_{1n}b_1+A_{2n}b_2+\cdots+A_{nn}b_n\end{bmatrix} \tag{3}$$

参照拉普拉斯展开式(P-I-2(3)式)，上式各行可以写成

$$x_i=\frac{1}{\det\boldsymbol{A}}\det\begin{bmatrix} a_{11} & \cdots & a_{1i-1}, & b_1, & a_{1i+1}, & \cdots, & a_{1n}\\ \vdots & & \vdots & \vdots & \vdots & & \\ a_{n1} & \cdots & a_{ni-1}, & b_n, & a_{ni+1}, & \cdots, & a_{nn}\end{bmatrix} \tag{4}$$

$$i=1,\ 2,\ \cdots,\ n$$

(4)式称为克兰姆公式(Cramer's rule).

注意，(4)式右边的矩阵是将 $\boldsymbol{A}$ 的第 i 列用 $\boldsymbol{b}$ 置换后得到的，而且由(4)式显而易见，解 $x_i\ (i=1, 2, \cdots, n)$ 是以和 a_{ij}, $b_i\ (i=1, 2, \cdots, n,\ j=1, 2, \cdots, n)$ 具有相同体 F 上的元给出.

其次，(1)式也包含 $\det\boldsymbol{A}=0$ 的情况. 在一般情况下，未知数的个数和方程式的个数不一定相等，即方程式包含 n 个未知数 $x_1, x_2, \cdots, x_n$，其系数是体 F 上的元的 m 个一次联立方程式

$$\left.\begin{array}{l}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2\\ \qquad\vdots\\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m\end{array}\right\}. \tag{5a}$$

现在我们来讨论这种方程式的解法. 先讨论方程式的一般性质, 即 (5) 式是否有解, 再进一步研究有几个解等问题.

由于

$$\boldsymbol{A}\triangleq\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{bmatrix}\triangleq[\boldsymbol{a}_1\boldsymbol{a}_2\cdots\boldsymbol{a}_n]\in F^{m\times n}$$

$$\boldsymbol{a}_j\triangleq \boldsymbol{A}\text{ 的第 } j \text{ 列 }\in F^m,\ \boldsymbol{b}\triangleq\begin{bmatrix}b_1\\ b_2\\ \vdots\\ b_m\end{bmatrix}\in F^m,\ \boldsymbol{x}\triangleq\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{bmatrix},$$

则(5)式可写成

$$\boldsymbol{Ax}=\boldsymbol{b} \tag{5b}$$

或

$$\boldsymbol{a}_1x_1+\boldsymbol{a}_2x_2+\cdots+\boldsymbol{a}_nx_n=\boldsymbol{b} \tag{5c}$$

当且仅当向量 $\boldsymbol{b}$ 可以用 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ 的线性组合表示时, (5c) 式的解才存在. 应用 B-I-3 定理 2, 可以将该条件叙述如下.

〔定理 1〕 当且仅当

$$\operatorname{rank}[\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n]=\operatorname{rank}[\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n|\boldsymbol{b}] \tag{6}$$

时, (5)式才有解存在.

因此, 设 $\operatorname{rank}\boldsymbol{A}=\operatorname{rank}[\boldsymbol{A}|\boldsymbol{b}]=\boldsymbol{r}$, 对于 $\boldsymbol{A}$ 可以应用 B-I-3 定理 3. 在方程式

$$\boldsymbol{Ax}=\boldsymbol{b}$$

两边左乘以 B-I-3 定理 3 中用过的 $\boldsymbol{P}_m$, 得

$$\boldsymbol{P}_m\boldsymbol{Ax}=\boldsymbol{P}_m\boldsymbol{b} \tag{7}$$

因 $\boldsymbol{P}_m$ 是正则矩阵, 则在方程式(5b) 和 (7) 当中, 其中一个方程式的解必定是另一个方程式的解. 在这个意义上, (5b)和(7)式等价.

其次, 利用 B-I-3 定理 3 中用过的 $\boldsymbol{P}_m$, 可以将(7)式变成下列形式

$$\boldsymbol{P}_m\boldsymbol{AP}_n\boldsymbol{P}_n^{-1}\boldsymbol{x}=\boldsymbol{P}_m\boldsymbol{b},$$

因为顺列矩阵是正交矩阵(P-I-1 问题 6), 即

$$\boldsymbol{P}_n^{-1}=\boldsymbol{P}_n^T$$

则

$$\boldsymbol{P}_m\boldsymbol{AP}_n\boldsymbol{P}_n^T\boldsymbol{x}=\boldsymbol{P}_m\boldsymbol{b}$$

式中若设

$$\boldsymbol{P}_m\boldsymbol{AP}_n\triangleq\tilde{\boldsymbol{A}},\ \boldsymbol{P}_n^T\boldsymbol{x}\triangleq\tilde{\boldsymbol{x}},\ \boldsymbol{P}_m\boldsymbol{b}\triangleq\tilde{\boldsymbol{b}}$$

则

$$\tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}}=\tilde{\boldsymbol{b}} \tag{8}$$

式中 $\tilde{\boldsymbol{x}}$ 是对 $\boldsymbol{x}$ 左乘以顺列矩阵 $\boldsymbol{P}_n^T$ 后得到的, 即在 $\boldsymbol{x}$ 中交换行, 结果使未知数的顺序更换了.

那么，若将(8)式中的 $\tilde{A}$, $\tilde{b}$, $\tilde{x}$ 分割成

$$\tilde{A}=\begin{array}{c} \begin{array}{cc} \overbrace{\phantom{\tilde{A}_{11}}}^{r} & \overbrace{\phantom{\tilde{A}_{12}}}^{n-r} \end{array} \\ \left[\begin{array}{c|c} \tilde{A}_{11} & \tilde{A}_{12} \\ \hline \tilde{A}_{21} & \tilde{A}_{22} \end{array}\right] \end{array}\begin{array}{l} \}r, \\ \}m-r \end{array}$$

$$\tilde{b}=\left[\begin{array}{c} b_1 \\ \hline b_2 \end{array}\right]\begin{array}{l}\}r\\ \}m-r\end{array}\quad \tilde{x}=\left[\begin{array}{c} \tilde{x}_1 \\ \hline \tilde{x}_2 \end{array}\right]\begin{array}{l}\}r\\ \}m-r\end{array} \tag{9}$$

由 B-I-3 定理 3 可知，$\tilde{A}_{11}$ 是 $r\times r$ 阶正则矩阵，$\tilde{A}_{22}=\tilde{A}_{21}\tilde{A}_{11}^{-1}\tilde{A}_{12}$. 根据假定，$\operatorname{rank}[A|b]=\operatorname{rank}[\tilde{A}|\tilde{b}]=r$，且因 $[\tilde{A}|\tilde{b}]$ 的第 1, 2, …, r 行是线性独立的，所以利用适当的基本变换可以将其变换成

$$[\tilde{A}|\tilde{b}]=\left[\begin{array}{c|c|c} \tilde{A}_{11} & \tilde{A}_{12} & \tilde{b}_1 \\ \hline \tilde{A}_{21} & \tilde{A}_{22} & \tilde{b}_2 \end{array}\right]\to\left[\begin{array}{c|c|c} \tilde{A}_{11} & \tilde{A}_{12} & \tilde{b}_1 \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} \end{array}\right] \tag{10}$$

事实上很容易看到，行变换矩阵是 $m\times m$ 阶正则矩阵

$$P_r=\begin{bmatrix} I_r & \mathbf{0} \\ \tilde{A}_{21}\tilde{A}_{11}^{-1} & I_{m-r} \end{bmatrix}.$$

因为 P_r 是正则矩阵，则在一个方程式的解必定是另一个方程式的解的意义上，方程式

$$\tilde{A}\tilde{x}=\tilde{b} \tag{8}$$

和

$$P_r\tilde{A}\tilde{x}=P_r\tilde{b} \tag{11}$$

等价. 因此，代替(8)式我们来讨论方程式(11). (11)式可写成

$$\tilde{A}_{11}\tilde{x}_1+\tilde{A}_{12}\tilde{x}_2=\tilde{b}_1 \tag{12}$$

任意假定一个 $\tilde{x}_2\in F^{n-r}$，把它当作已知向量，解 $\tilde{x}_1$ 得

$$\tilde{x}_1=\tilde{A}_{11}^{-1}(\tilde{b}_1-\tilde{A}_{12}\tilde{x}_2) \tag{13}$$

即解的一般形式为

$$\tilde{x}=\begin{bmatrix} \tilde{A}_{11}^{-1}(\tilde{b}_1-\tilde{A}_{12}\tilde{x}_2) \\ \tilde{x}_2 \end{bmatrix} \tag{14}$$

因此，若改为原来的未知数 x，则得到

$$x=P_n\tilde{x}=P_n\begin{bmatrix} \tilde{A}_{11}^{-1}(\tilde{b}_1-\tilde{A}_{12}\tilde{x}_2) \\ \tilde{x}_2 \end{bmatrix} \tag{15}$$

归纳以上得到如下定理.

[定理 2]　当 $\operatorname{rank}A=\operatorname{rank}[A|b]=r$ 时，对于如 (5) 式所示的线性方程组 $Ax=b$，适当选择$(n-r)$个未知数(与上述的 $\tilde{x}_2$ 对应)，其值可任意给定，但一旦给出后，其余 r 个未知数(对应于 $\tilde{x}_1$)的解便唯一确定.

(问题 1)　以上是在 B-I-3 定理 3 的基础上进行讨论，当然使用 A 的标准形也可以进行同样的考察. 设 A 的标准形为 $\tilde{A}_r(=PA)$，因矩阵 P 是正则矩阵，则

$$Ax=b$$

和

$$PAx=Pb$$

等价　因此，设 $\tilde{b}\triangleq Pb$，代替(5a)式可来解下列方程式

$$\tilde{A}_r x=\tilde{b},$$

$\{k_1, \cdots, k_r\}$ 按 B-I-2(5)式定义，设在 $\{1, \cdots, n\}$ 中除掉 $\{k_1, \cdots, k_r\}$ 后剩下的用符号 $\{k_1', \cdots, k_{n-r}'\}$ 表示. 试证明$(n-r)$个未知数 $x_{k_1'}, \cdots, x_{k_{n-r}'}$ 的值可任意给定，但其一旦给定后，余下 r 个未知数 $x_{k_1}, \cdots, x_{k_r}$ 的解是唯一确定的.

以上，是对方程式 (5) 有解的条件及其解的性质进行的一般性讨论，实际上是判定解的存在性. 为了对解进行数值计算，必须按上述原理拟定出一套计算方法. 关于线性方程组，有各种各样的数值解法[24]，这里不可能都加以说明，但是对于大家所熟悉的高斯消去法[25~27a](Gauss elimination method)和克劳特(Crout)法[25~27a]等要介绍一下. 特别是克劳特法，我们知道，在系数矩阵 $\boldsymbol{A}$ 中含有的多数是 0 元素(所谓的虫食矩阵)的情况下特别有效.

其次，我们来研究 $\boldsymbol{b}=\mathbf{0}$ 时 (5) 式的特殊情况，$\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$. 应用定理 1, 2 可以得到以下结果，求 $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$ 的解与在第 II 部分中求线性变换的零空间有直接关系，在系统理论中具有重要作用.

〔定理 3〕 设有系数矩阵为 $\boldsymbol{A}\in F^{m\times n}$，未知数为 $x_1, x_2, \cdots, x_n$ 的线性方程组$\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$，若 $\mathrm{rank}\boldsymbol{A}=r$，则具有 $(n-r)$ 个线性独立解，且方程式的任何解都可以用其线性组合表示.

(证明) 对应于 $\mathrm{rank}\boldsymbol{A}=r$，和定理 2 的情况一样，将 $\boldsymbol{A}$ 变换成 $\widetilde{\boldsymbol{A}}=\boldsymbol{P}_m\boldsymbol{A}\boldsymbol{P}_n$，当分割成

$$\widetilde{\boldsymbol{A}}=\left[\begin{array}{c|c}\widetilde{\boldsymbol{A}}_{11} & \widetilde{\boldsymbol{A}}_{12}\\ \hline \widetilde{\boldsymbol{A}}_{21} & \widetilde{\boldsymbol{A}}_{22}\end{array}\right]$$

时，$\widetilde{\boldsymbol{A}}_{11}=r\times r$ 是正则矩阵.

参照 (15) 式，因为 $\boldsymbol{b}=0$，当 $\tilde{\boldsymbol{x}}_2\in F^{n-r}$ 作为任意向量时，$\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$ 的一般解可以表示成

$$\boldsymbol{x}=\boldsymbol{P}_n\begin{bmatrix}-\widetilde{\boldsymbol{A}}_{11}^{-1}\widetilde{\boldsymbol{A}}_{12}\tilde{\boldsymbol{x}}_2\\ \tilde{\boldsymbol{x}}_2\end{bmatrix} \tag{16}$$

的形式. 在 (16) 式中，$\tilde{\boldsymbol{x}}_2$ 可选取 $(n-r)$ 个线性独立的值，例如在最简单的情况下选取 $\tilde{\boldsymbol{x}}_2=\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_{n-r}$，便得到$(n-r)$个线性独立解 $\boldsymbol{x}$. 现来看 $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$ 的任意解，因其必须具有(16)式的形式，显然可以用其$(n-r)$个特解

$$\boldsymbol{P}_n\begin{bmatrix}-\widetilde{\boldsymbol{A}}_{11}^{-1}\widetilde{\boldsymbol{A}}_{12}\boldsymbol{e}_1\\ \boldsymbol{e}_1\end{bmatrix}\cdots\boldsymbol{P}_n\begin{bmatrix}-\widetilde{\boldsymbol{A}}_{11}^{-1}\widetilde{\boldsymbol{A}}_{12}\boldsymbol{e}_{n-r}\\ \boldsymbol{e}_{n-r}\end{bmatrix}$$

的线性组合表示.

定理 3 是方程式 $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$ 的一般性质的叙述. 实际上，为了求 $(n-r)$ 个线性独立解 $\boldsymbol{x}$，最好利用 $\boldsymbol{A}$ 的标准形将定理 3 的原理具体化.

求 $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$ 的线性独立解的方法

设 $\boldsymbol{A}\in F^{m\times n}, \mathrm{rank}\boldsymbol{A}=r$.

(i) 利用 B-I-2 中讲过的基本变换，将所给系数矩阵 $\boldsymbol{A}$ 变换成行同值的行标准形矩阵 $\widetilde{\boldsymbol{A}}_r$.

(ii) 将 $\boldsymbol{A}$ 变换成

$$\begin{array}{c} \quad\quad\quad\quad 第k_1列 \quad\quad 第k_2列 \quad 第k_r列 \\ \widetilde{A}_r=\begin{bmatrix} 0 & 0 & \cdots 1 & * & \cdots\cdots & 0 & \cdots\cdots & 0 & \cdots \\ 0 & 0 & \cdots\cdots & \cdots\cdots & 0 & 1 & *\cdots\cdots & 0 & \cdots \\ \vdots & \vdots & & & & & & & \\ 0 & 0 & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & 0 & 1 & * \\ 0 & 0 & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots & & 0 \\ \vdots & \vdots & & & & & & & \\ 0 & 0 & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots & & 0 \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}0\\0\\0\\0\end{matrix}}\right\}r \\ \left.\vphantom{\begin{matrix}0\\0\\0\end{matrix}}\right\}m-r \end{matrix} \end{array}$$

代替 $\boldsymbol{Ax}=\mathbf{0}$，我们来研究其等价方程式 $\widetilde{\boldsymbol{A}}_r\boldsymbol{x}=\mathbf{0}$，解 $(n-r)$ 次，求得 $(n-r)$ 个独立解．为此，设 K 是正整数 $k_1, k_2, \cdots, k_r$ 的集合，K^c 为集合 $\{1, 2, \cdots, n\}$ 上 K 的补集，即令

$K=\{k_1, k_2, \cdots, k_r\}$，$K^c=\{1, 2, \cdots, n\}-K=\{k'_1, k'_2, \cdots, k'_{n-r}\}$．对于 $\boldsymbol{x}=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$，我们来考察 $\boldsymbol{y}=\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$，设其分量按下式

$$y_i=\begin{cases} x_j,\ j\notin K^c & （即 j\in K） \\ 1,\ j=k'_k \\ 0,\ j\in K^c,\ j\neq k'_k \end{cases}$$

给出的向量为 $\boldsymbol{y}_k$，对于 $k=1, 2, \cdots, n-r$，总共得 $(n-r)$ 个 $\boldsymbol{y}_1, \cdots, \boldsymbol{y}_{n-r}$．在 $(n-r)$ 个方程式

$$\left.\begin{matrix} \widetilde{\boldsymbol{A}}_r\boldsymbol{y}_1=\mathbf{0} \\ \vdots \\ \widetilde{\boldsymbol{A}}_r\boldsymbol{y}_{n-r}=\mathbf{0} \end{matrix}\right\}$$

中，对于 $\boldsymbol{y}_1, \cdots, \boldsymbol{y}_{n-r}$ 中的未知分量求解．

(iii) 在 (ii) 中所确定的 $\boldsymbol{y}_1, \cdots, \boldsymbol{y}_{n-r}$ 是所要求的 $(n-r)$ 个线性独立解．

（例）

$$A=\begin{bmatrix} 1 & 3 & 2 & 1 \\ 2 & 0 & 1 & 2 \\ 3 & 1 & 2 & 3 \end{bmatrix}\in R^{3\times 4}$$

(i) $\boldsymbol{A}$ 的行等价标准形

$$\widetilde{\boldsymbol{A}}_r=\begin{bmatrix} 1 & 0 & \frac{1}{2} & 1 \\ 0 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

(ii) $n=4$, $k_1=1$, $k_2=2$, $r=2$，因此 $K=\{1, 2\}$

$$K^c=\{1, 2, 3, 4\}-\{1, 2\}=\{3, 4\}\triangleq\{k'_1, k'_2\}$$

因此，$\boldsymbol{Ax}=\mathbf{0}$ 的线性独立解有两个，

$$\boldsymbol{y}_1=\begin{bmatrix}x_1\\x_1\\1\\0\end{bmatrix}\begin{matrix}\left.\begin{matrix}\\ \\ \end{matrix}\right\}K\\ \left.\begin{matrix}\\ \\ \end{matrix}\right\}K^o\end{matrix}\qquad \boldsymbol{y}_2=\begin{bmatrix}x_1\\x_2\\0\\1\end{bmatrix}\begin{matrix}\left.\begin{matrix}\\ \\ \end{matrix}\right\}K\\ \left.\begin{matrix}\\ \\ \end{matrix}\right\}K^o\end{matrix}$$

$$\tilde{\boldsymbol{A}}_r\boldsymbol{y}_1=\boldsymbol{0}:\begin{bmatrix}1&0&\frac{1}{2}&1\\0&1&\frac{1}{2}&0\\0&0&0&0\end{bmatrix}\begin{bmatrix}x_1\\x_2\\1\\0\end{bmatrix}=\begin{bmatrix}0\\0\\0\\0\end{bmatrix}\Rightarrow\boldsymbol{y}_1=\begin{bmatrix}-\frac{1}{2}\\-\frac{1}{2}\\1\\0\end{bmatrix}$$

$$\tilde{\boldsymbol{A}}_r\boldsymbol{y}_2=\boldsymbol{0}:\begin{bmatrix}1&0&\frac{1}{2}&1\\0&1&\frac{1}{2}&0\\0&0&0&0\end{bmatrix}\begin{bmatrix}x_1\\x_2\\0\\1\end{bmatrix}=\begin{bmatrix}0\\0\\0\\0\end{bmatrix}\Rightarrow\boldsymbol{y}_2=\begin{bmatrix}-1\\0\\0\\1\end{bmatrix}$$

（问题 2） 在上述求 $\boldsymbol{Ax}=\boldsymbol{0}$ 的线性独立解的方法中，不一定非要使用 $\boldsymbol{A}$ 的标准形，试证明也可以使用 B-I-2 行标准化步骤(i)～(iv)所得到的阶梯形.

（问题 3） 根据 B-I-3 定理 4，将 $\boldsymbol{A}$ 分解成都具有最大秩矩阵的乘积，设为 $\boldsymbol{A}=\boldsymbol{BC}$，试证明 $\boldsymbol{Ax}=\boldsymbol{0}$ 的充分和必要条件是 $\boldsymbol{Cx}=\boldsymbol{0}$.

A-I-1　系统及其状态方程式

以上各章我们介绍了关于矩阵的各种基础知识，一定会有些读者要问，何时介绍矩阵在自动控制中的应用呢？因此从本章起，我们将从如何应用前面所讲的矩阵理论的角度来讨论系统及其状态方程式．具体地讲，对于所给的系统，讨论应该用什么方法确定出其状态，以及如何推导其状态方程式等问题．因为求解状态方程式要求具备矩阵解析的知识，所以那是以后要研究的内容．

在讲述推导系统状态方程式的方法之前，先解释一下系统理论中的几个术语，以免在以后的讨论中引起混淆．

系统模型及其表示[28]

系统这个术语对自动控制工作者是非常熟悉的，但是对许多术语的意义和用法若不注意，常常会发生混淆．例如，我们经常说"对于系统 $\dot{\boldsymbol{x}}=\boldsymbol{Ax}+\boldsymbol{Bu}$，$\boldsymbol{y}=\boldsymbol{Cx}+\boldsymbol{Du}$……"，但是这时我们的真正意思是说"对于系统的数学表示(或数学模型) $\dot{\boldsymbol{x}}=\boldsymbol{Ax}+\boldsymbol{Bu}$，$\boldsymbol{y}=\boldsymbol{Cx}+\boldsymbol{Du}$……"，而丝毫没有把实际系统直接当作对象的意思．更准确点讲，它不是能够完全地表示实际系统特性的方程式，而是指在各种假定下将实际系统加以简化后得到的系统模型，不过是用数学公式表示模型动态特性的一种形式而已．对大家来讲，把这一点搞清楚是非常重要的．于是必须明确，当我们谈到"系统"时是指下列概念中的那一个．

(i) 实际系统

顾名思义，它是我们完全能够看到、感触到或使用的系统．例如，听立体音乐的高保真度放大器，生产过程，卫星通讯系统等物理系统，以及作为城市规划对象的东京、大阪等等．

(ii) 系统模型

在工程上，大多用模型来代替实际系统进行设计、分析和讨论．所谓模型，是技术人员根据设计、分析的目的，只注意实际系统某一方面的特性，将实际系统进行简化而得到的．采用模型是技术人员常用的一种方法，这时按照技术人员的意图，对于一个实际系统，通常可以得出各种各样的模型．例如，利用热电偶和继电器组成的使恒温槽中温度保持一定的简单控制系统上，在研究温度分布的问题时，控制对象(即恒温槽)可以采用一个具有分布参数的模型；若只考察恒温槽中的平均温度，还可以将模型进一步简化成一个具有集中参数的一阶惯性系统．

在历磁一定的直流电机上，将电枢电压作为输入，转子转角作为输出，通常忽略转轴的库伦摩擦转矩，使模型简化成为一个转矩和电枢电流成比例的具有集中参数的线性系统．如果考虑库伦摩擦转矩和齿轮的啮合间隙，转矩和电机电枢电流饱和的非线性特性等，模型必然是非线性的．

(iii) 模型的数学表示

使用模型进行设计、分析时，如能将其模型用数学公式表示，当然是很方便的．但是

即使对于同一个模型，由于变量的选择方法以及观察特性的角度不同，也会有各种各样的表示形式.

（例1） 现来看图1所示的简单电路. 设图中端子 A, B 间的电压 u 为输入量，由端子 A 流入的电流 y 为输出量. 这是一个由 1Ω 电阻、1F 电容及 1H 电感组成的线性常非时变系统（即系统特性不随端子电压、电流及时间发生变化）. 它是将电阻、电容、电感理想化之后，由实际系统得到的模型. 在该电路中，若选取电容器上电压 x_1、电感中电流 x_2 为状态变量（参看图 1(a)），则得到

电容器中电流* $=-2x_1+x_2+u$

电感上电压 $=-x_1+u$,

上式左边分别等于 $\dot{x}_1$, $\dot{x}_2$, 代入后便得到下列状态方程式

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix}=\begin{bmatrix}-2 & 1\\ -1 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}+\begin{bmatrix}1\\ 1\end{bmatrix}u \tag{1a}$$

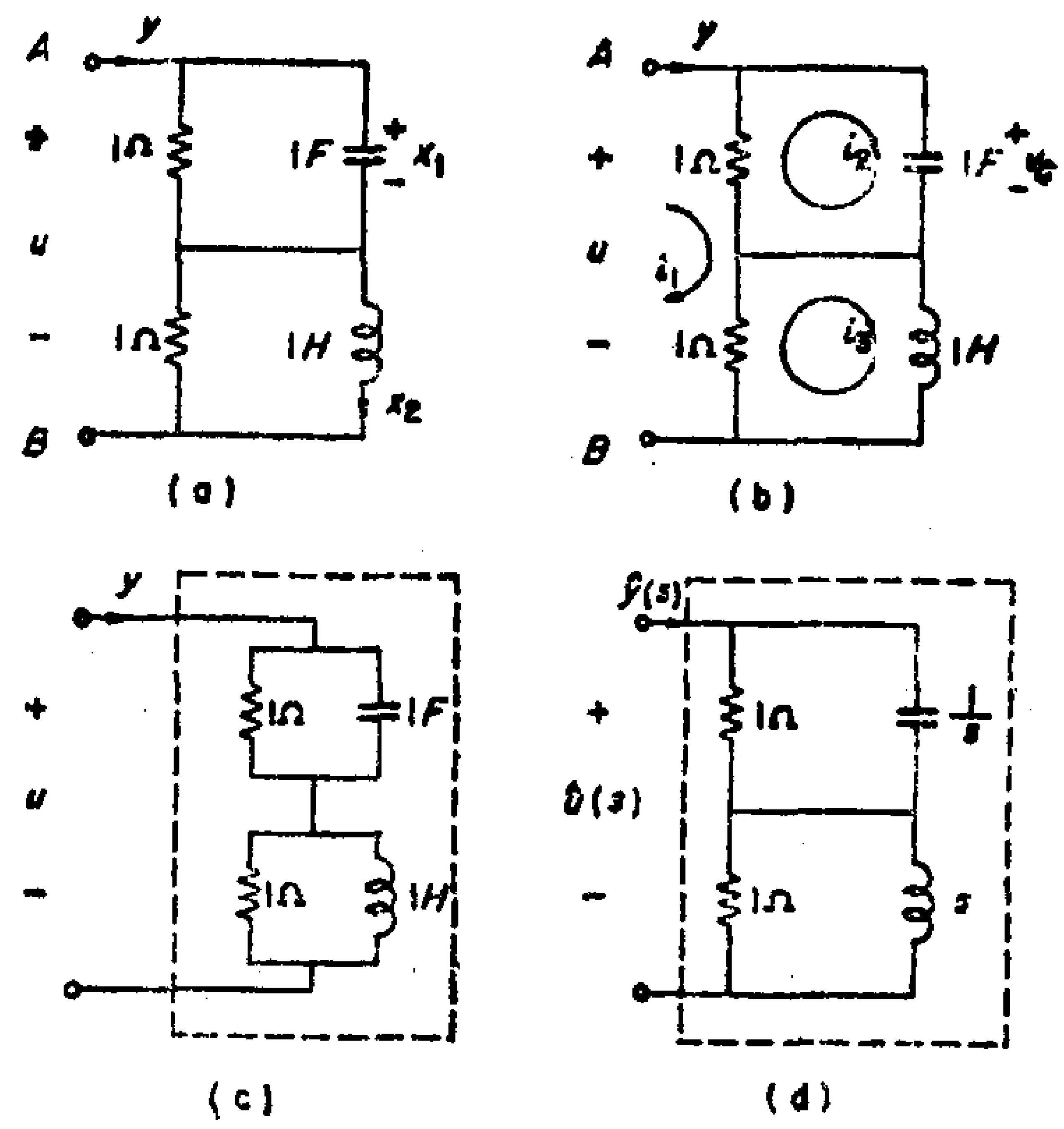


图1 线性非时变模型(电气回路)

根据克希霍夫第一定律，得

$$y=(\text{电感中电流})+(\text{下边 }1\Omega\text{ 电阻中的电流})^{**}=-x_1+x_2+u \tag{1b}$$

(1)式是表示该回路(模型)使用状态变量的动态特性.

另外，在观察回路电流 i_1、i_2、i_3 的情况下（图 1(b)），对各回路应用克希霍夫第二定律，便得到下列回路电压方程式

$$\begin{aligned}&2i_1-i_2-i_3=u\\ &-i_1+i_2+\int_0^t i_2(\tau)d\tau+v_c(0)=0\\ &-i_1+i_3+\dot{i}_3=0\end{aligned} \tag{2a}$$

式中 $v_c(0)$ 是 $t=0$ 时电容器上的电压，输出由式

$$y=i_1 \tag{2b}$$

给出. (2)式是按回路电流的回路表达式.

那么，若求出(1)式((2)式)的解，则该解就确定了回路中一切元件上的端电压、电流. 在这个意义上，(1)((2)式)可以称为含有回路内部信息的方程式. 例如，图 1(a) 中上边电阻中的电流是 x_1. 下边电阻中的电流为 $(-x_1+u)$，两者都由 x_1, x_2 确定. 在图 1(b) 中，它们分别以 (i_1-i_2), (i_1-i_3) 给出. 在许多情况下，对该回路我们只关心回路端子上的输入-输出响应，而不必要知道关于内部状态的信息. 在图 1(c) 中，从端子 A, B 观察该回路时，为了求出输入量 u 和输出量 y 之间的微分方程式（输入输出微分方程式），只要由(1)式(或(2)式)消去中间变量即可. 这时，将(1a)（或(1b)）式进行适当的微分，消去 x_1 和 x_2 便可得到 u 和 y 之间的微分方程式. 但是必须注意，不适当的微分将会导致 $u-y$ 之间微分方程式阶数的不必要增高.

* 原文误为电容器上电压. ——译者注

** $y=x_2+(u-x_1)=-x_1+x_2+u$ 原文为 $y=$(电容器中的电流)+(上边 1Ω 电阻中的电流)，这样推导既烦琐又不易理解. ——译者注

由(1)式或(2)式形式的方程式推导输入 u 和输出 y 之间的输入输出微分方程式的详细说明,涉及多项式元素的矩阵,在本书的第 III 部分中将要叙述,这里只要求大家注意,要推导出阶数尽可能低的输入输出微分方程式.在该例中可如下进行.将(1b)式微分,得

$$\dot{y}=\dot{x}_2-\dot{x}_1+\dot{u}$$

将(1a)式代入,消去 $\dot{x}_2$, $\dot{x}_1$, 再利用(1b)式,得

$$\dot{y}+y=\dot{u}+u \tag{3}$$

这就是该回路的输入输出微分方程式.

从端子看进去,除了用上面得到的微分方程式表示输入输出响应外,还有大家所最熟悉的传递函数.传递函数是初始条件(在该例中是 $t=0$ 时电容器上的电压、电感中的电流)全为 0 时输出量与输入量的拉普拉斯变换之比 $\hat{y}(s)/\hat{u}(s)$. 在该回路中,$\hat{y}(s)/\hat{u}(s)$ 是从端子 A, B 看进去的导纳.大家知道,它可以很简单地由图 1(d)所示等效电路求出

$$\text{传递函数 } \hat{h}(s)\triangleq\frac{\hat{y}(s)}{\hat{u}(s)}=1 \tag{4}$$

在零起始条件下,该回路由端子 A, B 看进去相当于 1 欧姆的电阻,并且不随时间变化.

最后,我们知道图 1 所示回路虽已用上述数学公式表示,但由于变量的选择及考虑问题的出发点不同,同一个回路可以有各种数学表示.在自动控制中,实际系统、模型和数学表示之间的差别是非常重要的.在控制理论及系统理论上,一般只是把模型和数学表示作为对象,但是由实际系统得到模型的过程,对自动控制工作者来说是非常重要而且困难的工作,也是充分发挥技术人员分析判断能力的地方.例如,如何建立人手的模型、一个国家经济的模型等等就可想而知了.

动力学系统和状态

本章的内容是讨论系统模型及其数学表示,不涉及到由实际系统得出模型的问题.而且,所讨论的模型仅限于所谓的动力学系统(dynamical system)和典型模型.这里我们也仅限于讨论由有限个具有集中参数的元件构成的系统模型,下面将介绍一下这种系统[1],并叙述其状态变量的定义.

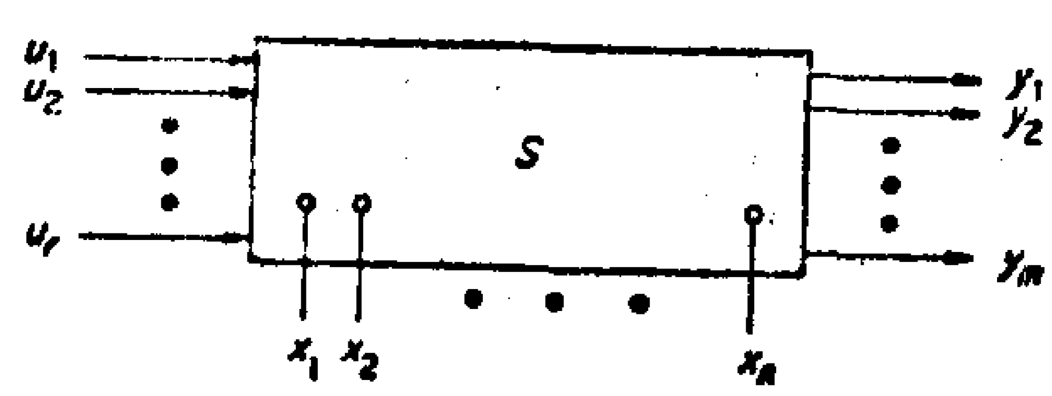


图 2 具有集中参数的系统模型

所谓动力学系统,是能够用状态变量的响应特性表示的系统模型.现在,我们来讨论具有 r 个输入量 u_1, u_2, ……, u_r 和 m 个输出量 y_1, y_2, ……, y_m, 由有限个具有集中参数的元件组成的系统模型 S(图 2).具体地讲,象例 1 中的回路等等.

所谓输入和输出,是表示系统与其外部相互作用的变量.即输入由系统外部(例如操作该系统的人)确定,它表示外部对系统的作用.相反,输出表示系统对外部的作用,具体的讲,例如是处于系统之外的人能够从系统得到的观测量.与此相反,下面定义的状态变量因与系统的内部状态有关,一般它们不是都能够由外部直接操作、直接观测的.

[1] 动力学系统是在控制理论、系统理论、回路理论、通信理论上经常用到的概念,它不仅包括具有集中参数的模型,而且还包括具有分布参数的更广泛的概念.例如,一般动力学系统的定义可参阅文献[28].

这里与模型 S 的响应特性有关，当我们能够找出同时具有下列性质(P.1)和(P.2)的有限个变量组 $x_1, x_2, \cdots\cdots, x_n$ 时，象这样的模型 S 称为有限维动力学系统，这些变量称为 S 的状态变量.

(P.1) 时刻 t 的状态变量值，由 $t_0 \leqslant t$ 时的起始状态变量值和从 t_0 到 t 的输入唯一确定;

(P.2) 时刻 t 的输出，由该时刻的状态变量值和输入唯一确定.

为了使符号简化，现在我们将输入和输出作为分量定义下列向量

$$\boldsymbol{u} \triangleq \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_r \end{bmatrix}, \quad \boldsymbol{y} \triangleq \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}, \tag{5}$$

若在时间 t 时输入 $u_i(t)$ $(i=1, 2, \cdots, r)$ 和输出 $y_j(t)$, $(j=1, 2, \cdots, m)$ 都是体 F 上的元，则 $\boldsymbol{u}(t) \in F^r$, $\boldsymbol{y}(t) \in F^m$(参看 B-0-2 中例 6). 在讲座的第 I, II, III 部分中，都假定 $u_i(t)$, $y_i(t)$ 在实数体 R 上取值；而在第 IV 部分中，则假定它们是有限体 $GF(p)$ 上的元(B-0-1 例 9).

同样，下列以状态变量为分量的向量

$$\boldsymbol{x} \triangleq \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{6}$$

称为状态向量(state vector). 时间 t 时状态向量 $\boldsymbol{x}(t)$ 所在的向量空间称为状态空间(state space). 控制理论中所遇到的状态空间主要是 R^n 或 C^n. 在讲座第 IV 部分所讨论的线性顺序系统中，因为状态变量是 $GF(p)$ 中的元，所以状态空间变成 $[GF(p)]^n$[1].

若利用输入、输出向量和状态向量，则因上述性质(P.1)与 $\boldsymbol{x}(t)$ 可作为 $\boldsymbol{x}(t_0)$ 和区间 $[t_0, t]$ 上输入向量 $\boldsymbol{u}[t_0, t]$ 的函数确定是一回事，故可用数学式表示如下[2]

(P.1)′
$$\boldsymbol{x}(t) = \boldsymbol{S}(t; \boldsymbol{x}(t_0), t_0; \boldsymbol{u}[t_0, t]), \quad t \geqslant t_0 \tag{7}$$

式中

$$\boldsymbol{u}[t_0, t] \triangleq [t_0, t] \text{上的输入}$$

同样，(P.2)也可表示成

(P.2)′
$$\boldsymbol{y}(t) = \boldsymbol{r}(t; \boldsymbol{x}(t); \boldsymbol{u}(t)) \tag{8}$$

(P.1)′和(P.2)′的内容可简述如下:“动力学系统的行为(状态和输出)由起始状态及其以后的输入完全确定”. 从几何学的角度来看(P.1)′意味着，在状态空间中从起始点

---

1) 体 F 为 $F=GF(p)$ 时，向量空间 F^n 写成 $[GF(p)]^n$.

2) 设 $x_i(t)$ 作为 t, $\boldsymbol{x}(t_0)$, t_0, $\boldsymbol{u}[t_0, t]$ 的函数可表示成

$$x_i(t) = S_i(t; \boldsymbol{x}(t_0), t_0; \boldsymbol{u}[t_0, t]) \quad 1 \leqslant i \leqslant n$$

以 $S_i(t; \boldsymbol{x}(t_0), t_0; \boldsymbol{u}[t_0, t])$ 为第 i 个元素的 $(n\times 1)$ 矩阵写成 $\boldsymbol{S}(t; \boldsymbol{x}(t_0), t_0; \boldsymbol{u}[t_0, t])$

$$\boldsymbol{S}(t; \boldsymbol{x}(t_0), t_0; \boldsymbol{u}[t_0, t]) \triangleq \begin{bmatrix} S_1(t; \boldsymbol{x}(t_0), t_0; \boldsymbol{u}[t_0, t]) \\ S_2(t; \boldsymbol{x}(t_0), t_0; \boldsymbol{u}[t_0, t]) \\ \vdots \\ S_n(t; \boldsymbol{x}(t_0), t_0; \boldsymbol{u}[t_0, t]) \end{bmatrix}.$$

(8)式的 $\boldsymbol{r}(t; \boldsymbol{x}(t); \boldsymbol{u}(t))$ 以及下面(10)式的 $\boldsymbol{f}(t; \boldsymbol{x}(t); \boldsymbol{u}(t))$ 等也是一样.

$\boldsymbol{x}(t_0)$开始，与输入相对应，时刻变动着的点 $\boldsymbol{x}(t)$ 所描绘的轨迹只有一条．该轨迹称为状态轨线(state trajectory)(图 3)．

(注意)

(i) 指定出输入、输出．对系统中所有的状态变量都存在着如 (P.1)′、(P.2)′ 所示的响应特性，这种系统才称为动力学系统．必须注意，在动力学系统定义中要同时指出输入和输出的这样一个特点．例如，在例 1 的回路中，若不指出输入和输出(即只给出这样一个由 L、R、C 组成的电路)，则不是动力学系统．图 1(a) 指定 A, B 两端子间的电压为输入，流入端子的电流为输出，这时响应特性可用方程式 (1a)，(1b)表示．如以后所述，能写成(1a)，(1b)就意味着(P.1)′，(P.2)′成立．因此，图 1(a) 所示电路是动力学系统．如果图中 x_1, x_2 照原样，输入和输出互相交换(端子电压作为输出，端子电流作为输入)，则得到不同的动力学系统．

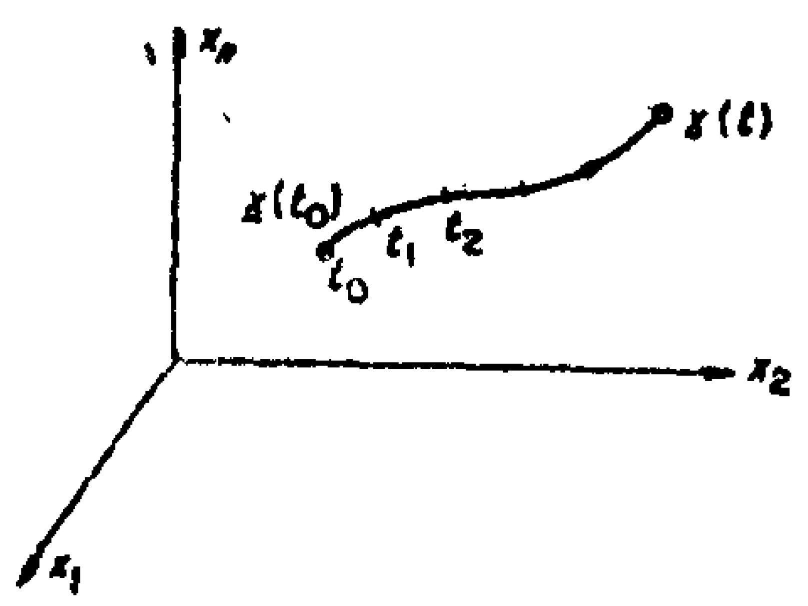


图 3 状态轨线

(ii) 在自动控制中遇到的系统模型大多是确定的动力学系统．但是，有时也把包含随机变量的系统作为对象．这里所定义的动力学系统没有考虑随机变量，因此从概率论的角度讨论系统时，必须扩大动力学系统概念的范畴．

(iii) 在(P.1)′、(P.2)′ 中，当研究在连续时间下动力学系统的行为时，认为 t(以及 t_0, $t\geqslant t_0$) 可取任意实数值，这种模型称为连续时间动力学系统 (continuous-time dynamical system)．

与此相反，在采样控制系统、顺序系统中等，是研究离散时刻 $t=t_0$, t_1, t_2, …… 下系统的行为．在这种情况下，认为(P.1)′，(P.2)′ 中的 t 取上面的离散时间值，这种模型叫做离散动力学系统(discrete-time dynamical system)．此外，在顺序系统中，时间在离散地变动之中，由于状态空间是$[GF(p)]^n$，所以时间 t 时状态向量 $\boldsymbol{x}(t)$仅可取有限个值(p^n)，在这个意义上，顺序系统等称为有限状态动力学系统(finite-state dynamical system)．

响应函数

将(P.1)′ 代入(P.2)′，输出也具有 (P.1)′ 同样的性质，即 $\boldsymbol{y}(t)$ 可看成由 $\boldsymbol{x}(t_0)$ 和 $\boldsymbol{u}[t_0, t]$决定的函数．即

$$\boldsymbol{y}(t)=\boldsymbol{r}(t;\boldsymbol{x}(t);\boldsymbol{u}(t))=\boldsymbol{r}(t;\overbrace{\boldsymbol{S}(t;\boldsymbol{x}(t_0),\ t_0;\boldsymbol{u}[t_0,\ t])}^{\boldsymbol{x}(t)};\boldsymbol{u}(t)),$$

将其改写成

$$\boldsymbol{y}(t)=\boldsymbol{\rho}(t;\boldsymbol{x}(t_0),\ t_0;\boldsymbol{u}[t_0,\ t]) \tag{9}$$

这时称(9)式右边的函数 $\boldsymbol{\rho}(\cdot;\cdot,\ \cdot;\cdot)$为动力学系统的响应函数(response function)[1]

状态方程式

这里我们来讨论对于给定模型找出状态变量的问题．

根据动力学系统的定义，求具有性质(P.1)，(P.2)的变量组 $(x_1, x_2, \cdots, x_n)$ 时，先假

[1] 如将变量 x 的函数写成 $f(x)$，容易和 x 下的 f 值相混淆，故表示成 $f(\cdot)$．

定其模型是动力学系统，并且$(x_1, x_2, \cdots, x_n)$是状态变量，最后看状态变量和动力学系统是否满足条件 (P.1)，(P.2)．但是这些条件本身并不能直接给出由模型求取状态变量的方法，(P.1) 是难于使用的条件．因此根据(P.1)，考虑下列实用的条件，它与 (P.1)(或(P.1)′)不等价(参看图 4)．

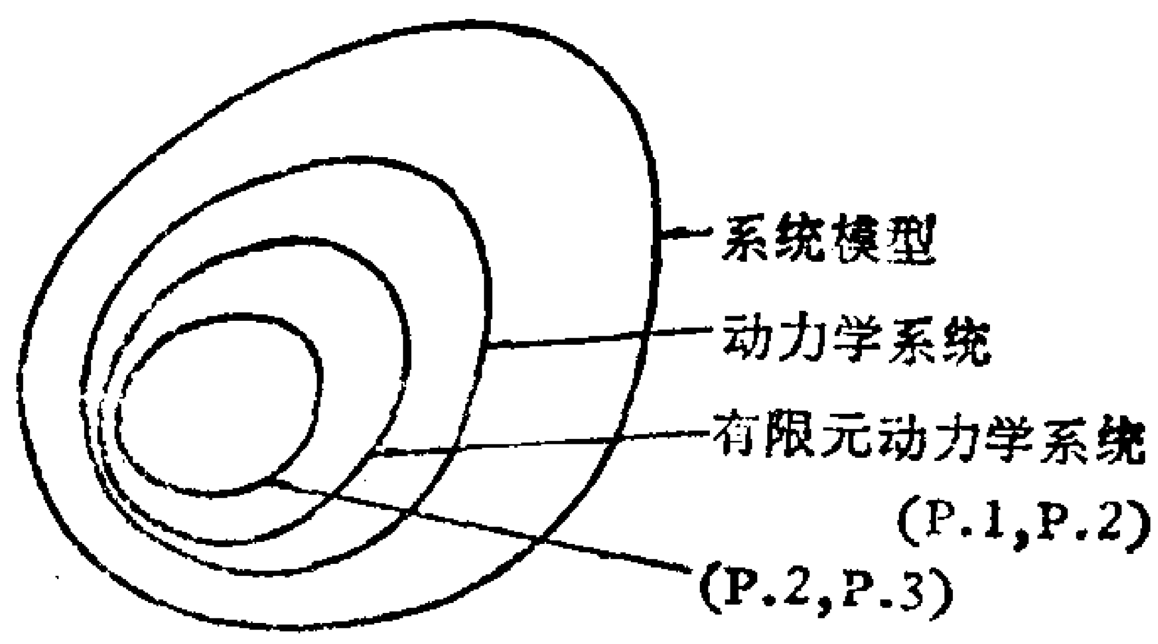


图 4　系统模型和动力学系统

(P.3) “对于连续时间模型，下列关于 $\boldsymbol{x}$ 的微分方程式

$$\dot{\boldsymbol{x}}(t)=\boldsymbol{f}(t;\boldsymbol{x}(t);\boldsymbol{u}(t)) \tag{10a}$$

成立．对于离散时间 $(t=t_0, t_1, t_2, \cdots)$ 模型，下列差分方程式

$$\boldsymbol{x}(t_{k+1})=\boldsymbol{f}(t_k;\boldsymbol{x}(t_k);\boldsymbol{u}(t_k)) \tag{10b}$$

成立．但是这里是假设对于任意初始状态方程式(10)的解都是唯一确定的”．

这里若详细地讲，“方程式(10)的解是唯一确定的”是指，当指定初始时间 t_0 下的状态 $\boldsymbol{x}(t_0)$，并且给出 t_0 以后的输入 $\boldsymbol{u}(t)$ $(t\geqslant t_0)$时，满足 $\boldsymbol{x}(t_0)=\boldsymbol{x}_0$ 的 (10) 式的解 $\boldsymbol{x}(t)$ 在 $t\geqslant t_0$ 时存在，而且只有一个．

由以上可见，(P.3)是(P.1)的充分条件．即在系统模型中若(P.3)成立，则保证(P.1)也成立．因此，对于给定动力学模型，为了找出状态变量 (而且表示模型也是动力学系统)，可按下列程序把(P.2)和(P.3)结合起来．

(i) 在模型中设定适当的变量$(x_1, x_2, \cdots, x_n)$，利用这些变量将输出用(8)式的形式表示出来(P.2)，其次推导出确定$(x_1, x_2, \cdots, x_n)$随时间变化的方程式，将其写成 (10) 式的形式，而且认为其解唯一存在(P.3)．

(ii) 这时$(x_1, x_2, \cdots, x_n)$就是模型的状态变量．

这里我们将对有限个元的系统求取状态变量的更一般的方法归纳成 (i) 和 (ii)．该法仅适用于能写成(10)式形式的动力学系统．大家知道，(10) 式是控制理论和系统理论中研究的对象，作为确定集中参数系统模型的基本表示方法被广泛采用．

本书只讨论动力学系统，一提到系统总是指动力学系统，而且是仅限于可用(8)，(10)式表示(即(P.2)，(P.3)成立)的动力学系统．

写成(10)式形式的微分(差分)方程式是一阶方程式，而 $\dot{\boldsymbol{x}}(\boldsymbol{x}(t_{k+1}))$是$t$, $\boldsymbol{x}(t)$, $\boldsymbol{u}(t)$ $(t_k, \boldsymbol{x}(t_k), \boldsymbol{u}(t_k))$的显函数．在这个意义上，称为正规形微分(差分)方程式(normal form differential (difference) equation)．一般说到状态方程式就是指这种正规形．

当(10)式右边是和输入及状态变量有关的线性函数，即

$$\dot{\boldsymbol{x}}(t)=\boldsymbol{A}(t)\boldsymbol{x}(t)+\boldsymbol{B}(t)\boldsymbol{u}(t) \tag{11a}$$

$$\boldsymbol{x}(t_{k+1})=\boldsymbol{A}(t_k)\boldsymbol{x}(t_k)+\boldsymbol{B}(t_k)\boldsymbol{u}(t_k) \tag{11b}$$

$$k=0,\ 1,\ 2,\ \cdots$$

(式中若 $\boldsymbol{x}(t)\in F^n$, $\boldsymbol{u}(t)\in F^r$, 则 $\boldsymbol{A}(t)\in F^{n\times n}$, $\boldsymbol{B}(t)\in F^{n\times r}$)时，该正规形方程式称为线性标准微分 (差分) 方程式，或简单地说成标准形微分 (差分) 方程式 (standard form differential (difference) equation)．

在本书中，状态方程式用(11)式的标准形式给出，输出方程式 $\boldsymbol{y}(t)=\boldsymbol{r}(t;\boldsymbol{x}(t);\boldsymbol{u}(t))$ 右边是 $\boldsymbol{x}(t)$, $\boldsymbol{u}(t)$有关的线性函数，即用

$$\boldsymbol{y}(t)=\boldsymbol{C}(t)\boldsymbol{x}(t)+\boldsymbol{D}(t)\boldsymbol{u}(t) \tag{12}$$

(式中若 $\boldsymbol{y}(t)\in F^m$, $\boldsymbol{x}(t)\in F^n$, $\boldsymbol{u}(t)\in F^r$, 则 $\boldsymbol{C}(t)\in F^{m\times n}$, $\boldsymbol{D}(t)\in F^{m\times r}$) 表示的系统称为线性系统.

(注意)

(i) 用正规形方程式(10)表示变量 $\boldsymbol{x}$, 并不能自动地保证其解的唯一存在(即对于 $\boldsymbol{x}$, (P.1) 不一定成立). 例如对于微分方程式 $\dot{x}=\sqrt{x}$, 当初始条件 $x(0)=0$ 时, 其解为 $x(t)=0$ 和 $1/4t^2$, 并不是唯一的. 因此,用 $\dot{x}=\sqrt{x}$ 表示的变量不是系统的状态变量. 在正规形方程式中,为了保证解的存在和唯一性, 作为附加条件, 对函数 $\boldsymbol{f}$ 的形式必须加以限制. 对于(10a)形式的微分方程式,以后要讲到里卜西兹条件是充分的[1]. 对于(10b)的差分方程式,有更低的条件,例如,若 $\boldsymbol{f}$ 的连续性成立就是充分的.

在线性的情况下(即在(11)式中), 如以后所述, 例如非常自然的条件是,当 $\boldsymbol{A}$, $\boldsymbol{B}$ 以 t 的分段连续的函数为元素时就是充分的. 因此, 标准形方程式成立大体上也可以说成是状态变量应该满足的条件(P.1)也同时成立的情况.

标准形状态方程式的推导方法概述

如上所述, 将所给系统用(10)式形式关于变量 x_1, x_2, $\cdots$, x_n 的正规形方程式表示非常重要,其理由是, (i) 以它为基础, 用理论方法或数值解法研究系统的动态特性时非常方便,(ii)它也间接证明 x_1, x_2, $\cdots$, x_n 是系统的状态变量.

关于线性系统,推导其标准形状态方程式的方法已很好研究过. 但是与此相反,对于一般的非线性系统,推导正规形方程式(10)的方法还没有很好地系统化(参考表 1).

表 1

标准形线性状态方程式 $\dot{\boldsymbol{x}}(t)=\boldsymbol{A}\boldsymbol{x}(t)+\boldsymbol{B}\boldsymbol{u}(t)$ 的推导方法
- 直接法
  - 由线性系统的图直接推导的方法
  - 由线性系统的方框图推导的方法
- 间接法
  - 由线性系统的其它数学表示(高阶微分方程组脉冲响应、传递函数)推导的方法
  - 利用非线性系统正规形方程式线性化的方法

若只限于标准形状态方程式 $\dot{\boldsymbol{x}}(t)=\boldsymbol{A}\boldsymbol{x}(t)+\boldsymbol{B}\boldsymbol{u}(t)$, 其推导方法可大致分为直接法和间接法两种. 前者推导时考虑系统各组成元素的性能及其之间的结合, 后者是从对系统应用物理法则或从观测系统响应得到其它数学表示求取标准形(参看表 1).

既使对于具有变系数矩阵的标准形 $\dot{\boldsymbol{x}}(t)=\boldsymbol{A}(t)\boldsymbol{x}(t)+\boldsymbol{B}(t)\boldsymbol{u}(t)$,上述常系数情况下的方法仍然适用. 下面在 A–I–2, A–I–3, A–I–4 将详细说明如何用图表示线性系统以及由方框图推导标准形的直接方法. 关于由高阶微分方程组、脉冲响应、传递函数等数学表示求取标准形的间接方法, 只进行一般性的说明. 因为它需要第 III 部分将要讲的以有理函数为元素的矩阵理论知识,所以在第 III 部分基础理论的应用中予以说明.

在系统本来是非线性、其特性可以用非线性正规形微分方程式(10a)表示的情况下,当考察在某个特定解附近的微小变动时(由于输入变动或初始状态变动引起的),将(10a)

[1] 因输入是已知函数,将其固定,则(10a)式右边 $\boldsymbol{f}(t;\boldsymbol{x};\boldsymbol{u}(t))$ 变成 x 和 t 的函数. 将它改写成 $\boldsymbol{f}(t;\boldsymbol{x};\boldsymbol{u}(t))\triangleq\tilde{\boldsymbol{f}}(t;\boldsymbol{x})$, 对于这种 $\dot{\boldsymbol{x}}=\tilde{\boldsymbol{f}}(t,\boldsymbol{x})$ 形式的微分方程式,里卜西兹条件是其解唯一存在的充分和必要条件. 如以后所述,当 $\dot{\boldsymbol{x}}=\tilde{\boldsymbol{f}}(t;\boldsymbol{x})$ 满足大范围的利普希茨条件时,不仅保证象(P.3)中所要求的 $t\geqslant t_0$ 时解 $\boldsymbol{x}(t)$ 的唯一性(当然也包括存在性),而且也确定了 t_0 以前的解也是唯一的.

式线性化，即用忽略掉高次项后的线性方程式来代替，这种线性表示式便是标准形(11a)的形式．关于这种用线性化手段推导标准形的方法，将在第 I 部分矩阵分析讲了以后再加以说明．

由于以上原因，这里不可能对表 1 中各种方法都详细加以说明，但是可以通过一个简单的例子对这些方法加以介绍(图 5)．

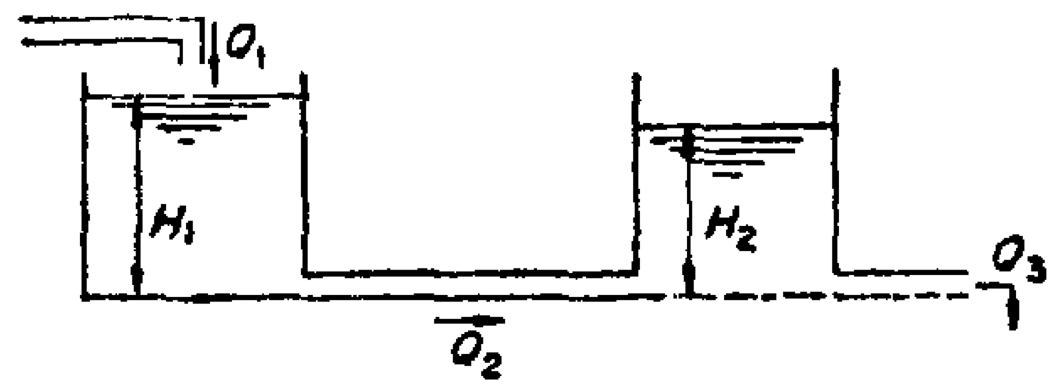


图 5　串联贮水箱系统

(例 2)[29] 现在我们来讨论截面积为 A_1，A_2 的两个水箱组成的串联贮水箱系统．设水箱 1 的流入量为 Q_1，水箱 2 的流入量为 Q_2，流出量为 Q_3，则当由水箱流出的过程呈现紊流状态时，Q_2 和水箱 1、2 中的水位差 H_1-H_2 的平方根成正比；Q_3 和 H_2 的平方根成正比．设其比例系数分别为 k_1，k_2，则可表示成

$$Q_2=k_1(H_1-H_2)^{1/2} \tag{13}$$

$$Q_3=k_2(H_2)^{1/2}, \tag{14}$$

水位 H_1，H_2 的变化分别由各水箱中流入量和流出量之差以下式给出

$$\frac{dH_1}{dt}=\frac{Q_1-Q_2}{A_1} \tag{15}$$

$$\frac{dH_2}{dt}=\frac{Q_2-Q_3}{A_2}, \tag{16}$$

由(13)～(16)式消去 Q_2，Q_3，得

$$\frac{dH_1}{dt}=\frac{-k_1}{A_1}(H_1-H_2)^{1/2}+\frac{Q_1}{A_1} \tag{17}$$

$$\frac{dH_2}{dt}=\frac{k_1}{A_2}(H_1-H_2)^{1/2}-\frac{k_2}{A_2}(H_2)^{1/2}, \tag{18}$$

这就是关于水位 H_1，H_2 的正规形微分方程式．

非线性微分方程式的线性化

现在，假定(17)，(18)式所示的串联贮水箱系统处于平衡状态，设这时水位和流量分别为 $H_1=\tilde{H}_1$，$H_2=\tilde{H}_2$，$Q_1=\tilde{Q}_1$，$Q_2=\tilde{Q}_2$，$Q_3=\tilde{Q}_3$，则下式成立，

$$\frac{-k_1}{A_1}(\tilde{H}_1-\tilde{H}_2)^{1/2}+\frac{\tilde{Q}_1}{A_1}=0 \tag{19}$$

$$\frac{k_1}{A_2}(\tilde{H}_1-\tilde{H}_2)^{1/2}-\frac{k_2}{A_2}(\tilde{H}_2)^{1/2}=0 \tag{20}$$

设其相对于平衡状态的变化量分别为 h_1，h_2，q_1，q_2，q_3，即

$$H_1\triangleq\tilde{H}_1+h_1,\quad H_2\triangleq\tilde{H}_2+h_2$$

$$Q\triangleq\tilde{Q}_1+q_1,\quad Q_2\triangleq\tilde{Q}_2+q_2,\quad Q_3\triangleq\tilde{Q}_3+q_3,$$

将(17)，(18)式在 $\tilde{H}_1$，$\tilde{H}_2$，$\tilde{Q}_1$ 的附近按泰勒级数展开，考虑到(19)，(20)，得

$$\frac{dh_1}{dt}=\frac{-1}{A_1R_1}(h_1-h_2)+\frac{q_1}{A_1}+(h_1,\ h_2\text{的高次项}) \tag{21a}$$

$$\frac{dh_2}{dt}=\frac{1}{A_2R_1}(h_1-h_2)-\frac{1}{A_2R_2}h_2+(h_1,\ h_2\text{ 的高次项}) \tag{21b}$$

式中　$$\frac{1}{R_1}\triangleq\frac{k_1}{2}(\tilde{H}_1-\tilde{H}_2)^{1/2},$$

$$\frac{1}{R_2} \triangleq \frac{k_2}{2}(\widetilde{H}_2)^{-1/2}.$$

(21)式是确定水位变化特性的微分方程式. 在这里若变化量非常微小, 略去 (21) 式右边的高次项仅保留一次项时, 可用下列线性化方程式

$$\begin{bmatrix} \dot{h}_1 \\ \dot{h}_2 \end{bmatrix} = \begin{bmatrix} -\dfrac{1}{A_1R_1} & \dfrac{1}{A_1R_1} \\ \dfrac{1}{A_2R_1} & -\left(\dfrac{1}{A_2R_1}+\dfrac{1}{A_2R_2}\right) \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} + \begin{bmatrix} \dfrac{1}{A_1} \\ 0 \end{bmatrix} q_1 \tag{22}$$

近似给出 h_1, h_2 的特性, 即(22)式是将正规形 (17)、(18) 式在平衡状态 $\widetilde{H}_1$, $\widetilde{H}_2$, $\widetilde{Q}_1$ 附近线性化后得到的标准形.

一般, 象(17), (18)式那样用下列

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t); \boldsymbol{u}(t)) \tag{23}$$

形正规形方程式表示的系统, 设在 $\boldsymbol{u}(t)=\boldsymbol{u}_0$, $\boldsymbol{x}(t)=\boldsymbol{x}_c$ 时处于平衡状态, 即

$$\boldsymbol{f}(\boldsymbol{x}_c, \boldsymbol{u}_0) = \boldsymbol{0}$$

成立, 对于该平衡状态 $(\boldsymbol{x}_c, \boldsymbol{u}_0)$ 的线性表示式象推导 (22) 式一样, 将 (23) 式右边在 $(\boldsymbol{x}_c, \boldsymbol{u}_0)$ 的附近用泰勒级数展开, 略去关于微小变化量 $\Delta \boldsymbol{x} \triangleq \boldsymbol{x}-\boldsymbol{x}_c$, $\Delta \boldsymbol{u} \triangleq \boldsymbol{u}-\boldsymbol{u}_0$ 的高次项, 可用下列标准形给出

$$\Delta\dot{\boldsymbol{x}} = \boldsymbol{A}\Delta\boldsymbol{x} + \boldsymbol{B}\Delta\boldsymbol{u} \tag{24}$$

式中系数矩阵各为

$$\boldsymbol{A} \triangleq \left[\frac{\partial f_i}{\partial x_j}\right]_{\substack{\boldsymbol{x}=\boldsymbol{x}_c \\ \boldsymbol{u}=\boldsymbol{u}_0}}, \quad \boldsymbol{B} \triangleq \left[\frac{\partial f_i}{\partial u_j}\right]_{\substack{\boldsymbol{x}=\boldsymbol{x}_c \\ \boldsymbol{u}=\boldsymbol{u}_0}}$$

其中 f_i 表示 $\boldsymbol{f}$ 的第 i 个元素, x_j 表示 $\boldsymbol{x}$ 的第 j 个元素.

以上是(23)式在平衡状态的线性化, 对于更一般情况下变化的输入 $\boldsymbol{u}(t)=\widetilde{\boldsymbol{u}}(t)$ 和相对它的解 $\boldsymbol{x}(t)=\widetilde{\boldsymbol{x}}(t)$, 线性表示式为

$$\Delta\dot{\boldsymbol{x}} = \boldsymbol{A}(t)\Delta\boldsymbol{x} + \boldsymbol{B}(t)\Delta\boldsymbol{u} \tag{25}$$

式中

$$\boldsymbol{A}(t) \triangleq \left[\frac{\partial f_i}{\partial x_j}\right]_{\substack{\boldsymbol{x}=\widetilde{\boldsymbol{x}}(t) \\ \boldsymbol{u}=\widetilde{\boldsymbol{u}}(t)}}, \quad \boldsymbol{B}(t) \triangleq \left[\frac{\partial f_i}{\partial x_j}\right]_{\substack{\boldsymbol{x}=\widetilde{\boldsymbol{x}}(t) \\ \boldsymbol{u}=\widetilde{\boldsymbol{u}}(t)}}.$$

根据(23)式, 即使对于更一般的正规形也是一样, 线性化表示式可以用(25)式形式的具有变系数矩阵的标准形给出.

上面表 1 中"按照非线性系统正规形方程式线性化的方法", 就是指用这样的过程推导出标准形.

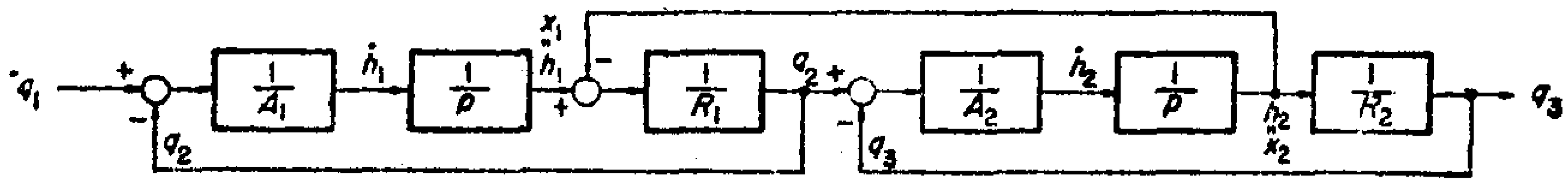


图 6 方框图

方框图

我们来看在起始平衡状态 $\widetilde{H}_1$, $\widetilde{H}_2$, $\widetilde{Q}_1$, $\widetilde{Q}_2$, $\widetilde{Q}_3$ 附近的微小变化量 h_1, h_2, q_1, q_2, q_3, 要分析其特性, 将(13)~(16)式线性化即可. 将(13), (14)式线性化后得

$$q_2 = \frac{h_1 - h_2}{R_1} \tag{26}$$

$$q_3=\frac{h_2}{R_2} \tag{27}$$

由此可见，流量的微小变化量与水位的微小变化量成正比．同样，由(15)，(16)式，

$$\frac{dh_1}{dt}=\frac{q_1-q_2}{A_1} \tag{28}$$

$$\frac{dh_2}{dt}=\frac{q_2-q_3}{A_2} \tag{29}$$

成立．用分析水箱 1，2 流入流出平衡的方法也可以推导出这两个表示式，(26)～(29) 式是在自动控制教科书中经常引用的双容水箱系统的线性模型关系式．

那么，在只考虑微小变化量的线性模型中，因模型各组成元素可以用 (26)～(29) 式表示，则 h_1，h_2，q_1，q_2，q_3 的关系可以表示在图 6 的方框图上．

方框 $1/p$ 表示一个积分器，设其输入量为 u，输出量为 y，则其输入输出特性可以用 $\dot{y}=u$ 表示．将输出 y 选为状态变量 x，即可写成标准形式

$$\begin{aligned}\dot{x}&=u\\ y&=x\end{aligned} \tag{30}$$

图中其它参数对应的方框是以相应参数为放大倍数的放大器．

现在，将各积分器的输出量（即 h_1 和 h_2）分别选为状态变量 x_1，x_2，考虑到各方框之间的输入输出关系，积分器的输入信号可以用 x_1，x_2，q_1 表示，积分器的输入信号不外乎是 $\dot{x}_1$，$\dot{x}_2$，将上述求得的值与 $\dot{x}_1$，$\dot{x}_2$ 相等，便得到下列标准形

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix}=\begin{bmatrix}-\dfrac{1}{A_1R_1} & \dfrac{1}{A_1R_1}\\ \dfrac{1}{A_2R_1} & -\left(\dfrac{1}{A_2R_1}+\dfrac{1}{A_2R_2}\right)\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}+\begin{bmatrix}\dfrac{1}{A_1}\\ 0\end{bmatrix}q_1 \tag{31}$$

输出方程式变成

$$q_3=\frac{x_2}{R_2} \tag{32}$$

这种由方框图求取标准形一般可以归纳成这样一个问题："在由用标准形 $\dot{\boldsymbol{x}}=\boldsymbol{Ax}+\boldsymbol{Bu}$，$\boldsymbol{y}=\boldsymbol{Cx}+\boldsymbol{Du}$ 表示的某些动态方框（在上例中相当于积分器方框）和放大器方框组成的方框图上，如何选择状态变量才能都用标准形表示"．该问题将在 A-I-4 讨论，根据方框图的构成用矩阵表示得到系统的计算方法．

微分方程组

在贮水箱系统的线性模型上，应用物理定律最后得到的(26)～(29)式，可以用矩阵写成

$$\begin{bmatrix}1 & 0 & -\dfrac{1}{R_1} & \dfrac{1}{R_1}\\ 0 & 1 & 0 & -\dfrac{1}{R_2}\\ \dfrac{1}{A_1} & 0 & p & 0\\ -\dfrac{1}{A_2} & \dfrac{1}{A_2} & 0 & p\end{bmatrix}\begin{bmatrix}q_2\\ q_3\\ h_1\\ h_2\end{bmatrix}=\begin{bmatrix}0\\ 0\\ \dfrac{1}{A_1}\\ 0\end{bmatrix}q_1 \tag{33}$$

式中 $p\triangleq\dfrac{d}{dt}$ 是微分算子．为了利用基本行变换将(33)式左边系数矩阵变换成分块上三角形矩阵，可进行如下操作：(i)将第1行×$(-1/A_1)$加到第3行；(ii)将第1行 ×$(1/A_1)$加到第4行；(iii)将第2行×$(-1/A_2)$加到第4行．如B-I-2所述，该结果和在(33)式左乘以 4×4 阶正则常数矩阵等价．最后，将(33)式变换成

$$\begin{bmatrix} 1 & 0 & -\dfrac{1}{R_1} & \dfrac{1}{R_1} \\ 0 & 1 & 0 & -\dfrac{1}{R_2} \\ 0 & 0 & \left(p+\dfrac{1}{A_1R_1}\right) & -\dfrac{1}{A_1R_1} \\ 0 & 0 & -\dfrac{1}{A_2R_1} & \left(p+\dfrac{1}{A_2R_1}+\dfrac{1}{A_2R_2}\right) \end{bmatrix}\begin{bmatrix} q_2 \\ q_3 \\ h_1 \\ h_2 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ \dfrac{1}{A_1} \\ 0 \end{bmatrix}q_1 \tag{34}$$

在一个方程式的解必定是另一个方程式解的意义上，(33)和(34)式是等价的．

整理(34)式的下面两行，便得到标准形(22)(或(31))式．刚才的问题可以一般地叙述如下，一般在不包含时变元素的具有集中参数的线性系统中，若能应用物理定律，便可得到下列形式的微分方程式

$$\begin{bmatrix} l_{11}(p)\cdots l_{1k}(p) \\ \vdots \\ l_{k1}(p)\cdots l_{kk}(p) \end{bmatrix}\begin{bmatrix} z_1 \\ \vdots \\ z_k \end{bmatrix}=\begin{bmatrix} m_{11}(p)\cdots m_{1r}(p) \\ \vdots \qquad\qquad \vdots \\ m_{k1}(p)\cdots m_{kr}(p) \end{bmatrix}\begin{bmatrix} u_1 \\ \vdots \\ u_r \end{bmatrix} \tag{35}$$

式中 $u_1, u_2, \cdots, u_r$ 是输入量，$z_1, z_2, \cdots, z_k$ 是为了表示系统而选择的物理变量，在上述例子中它相当于 q_2, q_3, h_1, h_2．(35)式的系数矩阵，一般是把以微分算子 $p\triangleq\dfrac{d}{dt}$ 的多项式作为元素的多项式环上的矩阵(参看B-0-1 例6)．(35)式在上例中相当于(33)式．在大家最熟悉的例子中象是分析电路时所使用的回路电流方程式、节点电位方程式都写成这种形式．由该形式的方程式推导标准形是一般性的问题，随后就要进行讨论．

传递函数

例如，在图6中将 p 换成 s，利用大家所熟悉的方框图简化方法即可求得该系统的传递函数

$$\frac{\mathscr{L}[q_3]}{\mathscr{L}[q_1]}=\frac{1}{A_1A_2R_1R_2s^2+(A_1R_1+A_2R_2+A_1R_2)s+1}. \tag{36}$$

(36)式传递函数不能唯一确定[1]表示该系统输入输出之间关系的微分方程式，但其中阶数最低的微分方程式是唯一确定的．在(36)式中用 p 代替 s，两边各乘以分母多项式，得

$$[A_1A_2R_1R_2p^2+(A_1R_1+A_2R_2+A_1R_2)p+1]q_3(t)=q_1(t), \tag{37}$$

式中 $$p\triangleq\frac{d}{dt}$$

将(37)式改写成标准形的最简单的方法是，令 $x_1\triangleq q_3$，$x_2\triangleq\dot{q}_3$ 为状态变量，于是得

[1] 例如，$(p+1)(p+2)y=(p+1)u$，$(p+1)(p+2)(p+3)y=(p+1)(p+3)u$．$(p+2)y=u$ 等等，都具有同一传递函数．

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \dfrac{-1}{A_1A_2R_1R_2} & \dfrac{-(A_1R_1+A_2R_2+A_1R_2)}{A_1A_2R_1R_2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \dfrac{1}{A_1A_2R_1R_2} \end{bmatrix} q_1$$

$$q_3 = x_1.$$

这个问题一般可叙述如下：

线性系统的传递函数(矩阵)$\hat{\boldsymbol{H}}(s)$，作为有理函数体(参看 B-0-1 例 7)上的矩阵，可以用下式给出，

$$\hat{\boldsymbol{H}}(s) = \begin{bmatrix} h_{11}(s) & h_{12}(s) & \cdots & h_{1r}(s) \\ h_{21}(s) & h_{22}(s) & \cdots & h_{2r}(s) \\ \vdots & & & \vdots \\ h_{m1}(s) & h_{m2}(s) & \cdots & h_{mr}(s) \end{bmatrix} \in Q(s,\ R)^{m\times r}.$$

具有这种传递函数的系统，其标准形方程式和输出方程式的求法是个问题，即使对于上面最简单的例子，具有同一传递函数 $\hat{\boldsymbol{H}}(s)$ 的标准形 $\dot{\boldsymbol{x}} = \boldsymbol{Ax} + \boldsymbol{Bu}$ 和输出方程式 $\boldsymbol{y} = \boldsymbol{Cx} + \boldsymbol{Du}$ 也不是唯一确定的，一般有无数个．找出其中具有最低维数的状态向量在应用上非常重要，在理论上也是很有意义的问题．

图*(graph)表示

图 5 所示系统用图表示非常方便，现将其改绘如下(图 7)．

图中 f 是接到水箱 1 的流量源，E 是具有大气压力 P_0 的水源，P_1，P_2 分别表示水箱 1，2 底部的水压．该水箱系统的图如图 8(a)所示．这里我们再来研究一下考虑压力、流量微小变化量下的线性模型．设其平衡状态值分别为 $\tilde{P}_1$，$\tilde{P}_2$，$\tilde{Q}_1$，$\tilde{Q}_2$，$\tilde{Q}_3$，则得变化量

$$p_1 \triangleq P_1 - \tilde{P}_1,\ p_2 \triangleq P_2 - \tilde{P}_2,$$
$$q_1 \triangleq Q_1 - \tilde{Q}_1,\ q_2 \triangleq Q_2 - \tilde{Q}_2,\ q_3 \triangleq Q_3 - \tilde{Q}_3,$$

在这里代替水位用压力作为变量，压力直接对应于电路中的电压，为了对比及用图表示方便，将电路及该例中的流体力学系统采取统一的表示形式．

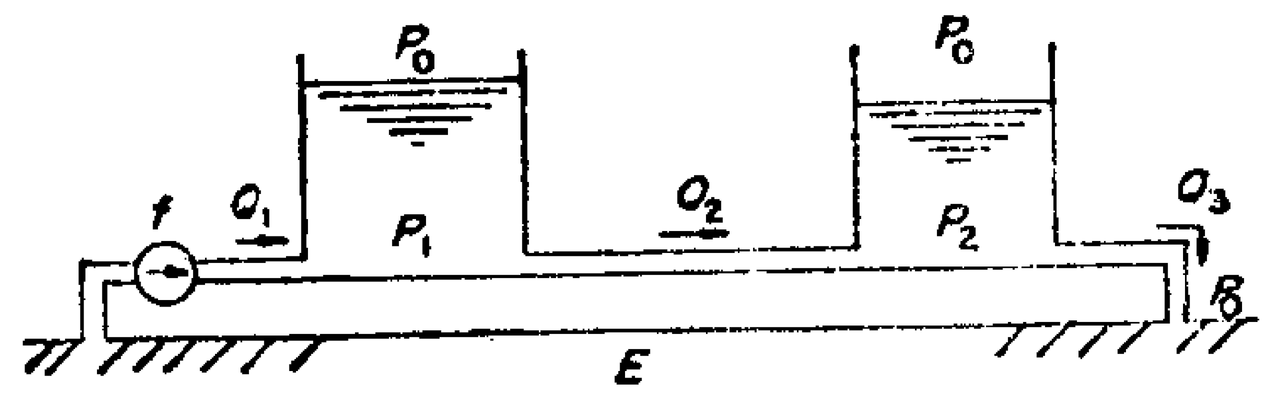


图 7　水源和贮水箱系统

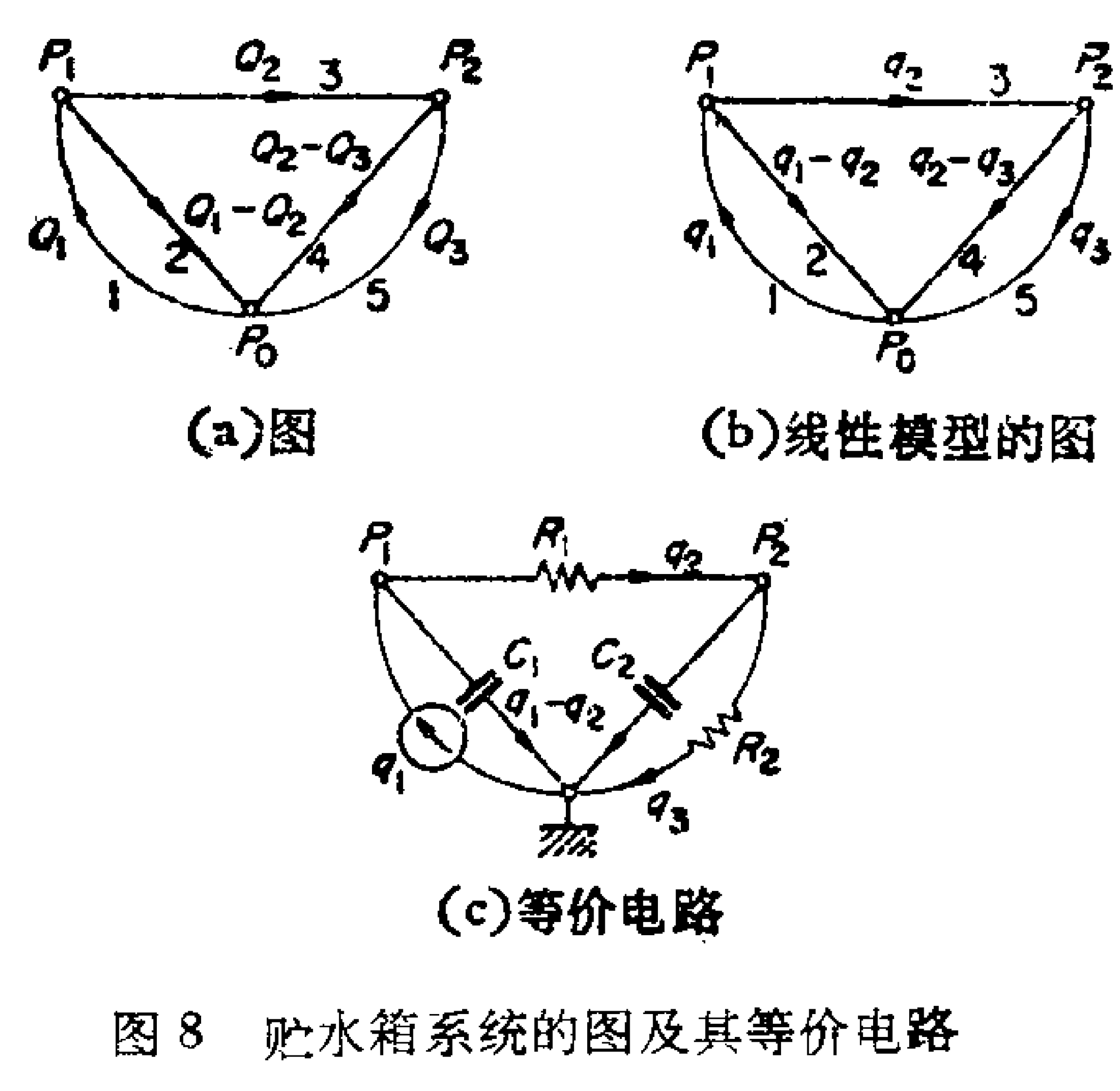


图 8　贮水箱系统的图及其等价电路

---

* 图论(theory of graph)起始研究 Königsberg 桥的问题，它是在拓扑学的基础上发展起来的．近年来在电气网络分析、设计及系统理论的研究中获得极为重要的应用．——译者注

那么图 8(b) 便是表示其变化量之间关系的图，图中考虑到 $p_1=h_1d$, $p_2=h_2d$($d\triangleq$ 水的比重$=1$)，根据(28)，(29)式，枝 2, 4 的特性(即水箱 1, 2 的特性)可以表示成

$$q_1-q_2=C_1\frac{dp_1}{dt},\ C_1\triangleq A_1^{*} \tag{37}$$

$$q_2-q_3=C_2\frac{dp_2}{dt},\ C_2\triangleq A_2^{**} \tag{38}$$

由(26)，(27)式可知

$$q_1=\frac{p_1-p_2}{R_1},\quad q_2=\frac{p_3}{R_2},$$

则枝 3, 5 (水管)具有图 8(b) 所示的图．对应枝用同一特性式表示的等价电路如图 8(c) 所示．因此象例 1 中一样，在该水箱系统的线性化模型上，把相当于图 8(c) 中 C_1, C_2 上电压 p_1, p_2 选为状态变量时，可以很容易地推导出其标准形．

这样，系统各组成元素之间的关系可以用图加以说明，运用在电气网络理论上发展起来的图论，可以由图推导其状态方程式．关于这方面的内容，将在 A-I-2, A-I-3 中继续详细讨论．

* 原文误为 $C_1\triangleq\frac{1}{A_1}$. ——译者注

** 原文误为 $C_2\triangleq\frac{1}{A_2}$. ——译者注

A-I-2 线性电气网络状态方程式的推导方法

本章内容

本章将详细研究，对于电气网络如何选择状态变量，用什么方法推导状态方程式的标准形(在其存在的情况下)．虽然是以具有非时变集中参数的元件构成的网络为主要对象，但是以能推广到包含有时变元件的线性时变网络的形式进行讨论，并且通过例题对各种情况下的结论加以说明．

本书之所以花这么多篇幅讨论电气网络，其理由如下：

(i) 因为在具有集中参数线性模型的机械系统、流体力学系统、热力学系统、电气-力学系统和电气网络之间有着显著的相似性，即和电气网络一样，其构成可以用图表示，其组成元素(适当选择变量)和电阻、电容、电感等有着相似的特性，所以对于电气网络推导状态方程式的方法，多半都可以原封不动地用于其它物理系统．

(ii) 大部分自动控制工作者对于电气网络比对其它物理系统要熟悉得多，尽量用它说明要方便些．

(iii) 电气网络可以立即表示成图，在图论的基础上进行分析，特别是以状态方程式为基础的解析法现在已得到很大发展[1)]．

本章分为两个部分，前部分讨论所谓的非时变 RLCM 网络状态方程式的推导方法[2)]，该网络由非时变的元件 R(电阻)、L(自感)、C(电容)、M(互感)组成．这部分得到的结论是，在不包含仅由电容组成的回路或仅由电感组成的割集(cut set*)(其意义以后说明)的普通结构的 RLCM 网络中，常常可以将电容器上的电压和电感中的电流选为状态变量，可以推导出标准形状态方程式．在这里可以得到关于电气网络状态方程式推导方法的一般知识，于是在此基础上就可以进行其它物理系统的讨论(**A-I-3**)．

本章后部分以包含下列电路元件的非时变(不随时间变化的)线性电气网络为对象，在这部分电路元件中，除了前面讨论的无源元件 R、L、C、M 以外，还有紧耦合电感、理想变压器、回转器，以及有源元件负电阻、从属电源．象这样一般的情况下，状态变量的选择是个问题，而且常常不能保证可以推导出状态方程式．因为必须考虑元件参数以及网络组成等各种复杂情况，所以不能得出一般性结论．在这种情况下，前面讲过的矩阵理论非常有用，特别是基本变换发挥了威力，它可以用来整理推导状态方程式的程序，使计算过程规律化．在这里将介绍这种适合于计算机的方法．

基本思考方法[23,31,32]

这里，我们研究一个简单电路状态方程式的推导方法．在图 1 电路中，输入是电流源

---

1) 由于利用计算机进行集成电路设计[CAD(Computer Aided Design)]的需要，使得分析电气网络的解析法得到很大发展[30]．

2) 假定 R、L、C、M 均为无源元件，即 $R>0$，$L>0$，$C>0$．对于互感，$L_1L_2-M^2>0$ 成立．

* 括号内英文系．——译者注

的电流 i_s，输出为电流源的端电压 v_s. 对应于该输入、输出，应如何选择满足条件(P.1)、(P.2)的状态变量？我们由经验知道，如果起始条件给成 $t=0$ 时电容器上的电压 v_c，电感中的电流 i_L，则 $t\geqslant 0$ 时电路的性能[包括 $v_c(t)$，$i_L(t)$，一切元件的电压、电流]是唯一确定的. 用公式形式可以写成

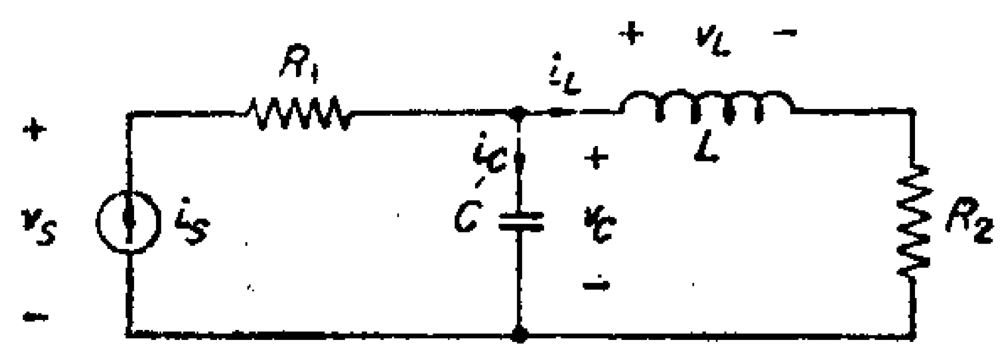


图1　简单的 RLC 电路

$$\begin{bmatrix} v_c(t) \\ i_L(t) \end{bmatrix}=\begin{bmatrix} s_1(t;\ v_c(0),\ i_L(0),\ 0;\ i_s[0,\ t] \\ s_2(t;\ v_c(0),\ i_L(0),\ 0;\ i_s[0,\ t] \end{bmatrix} \tag{1}$$

而且易见，输出可以写成

$$v_s(t)=v_c(t)+R_1 i_s(t) \tag{2}$$

将式(1)，(2)和条件(P.1)′，(P.2)′对照可见，在该电路中状态变量是我们预料中的$(v_c,\ i_L)$. 实际上，只要将关于$(v_c,\ i_L)$的微分方程式变成下列标准形式

$$\begin{bmatrix} \dot{v}_c(t) \\ \dot{i}_L(t) \end{bmatrix}=\boldsymbol{A}\begin{bmatrix} v_c(t) \\ i_L(t) \end{bmatrix}+\boldsymbol{b}i_s(t) \tag{3}$$

即可证明该结论的正确性. 由(3)式可见，如能将 $\dot{v}_c$、$\dot{i}_L$ 用 v_c、i_L 及 i_s 表示，即可得到标准形微分方程式. 因

$$\dot{v}_c(t)=\frac{1}{C}i_c(t),\quad \dot{i}_L(t)=\frac{1}{L}v_L(t)$$

则 $\dot{v}_c$，$\dot{i}_L$ 用 v_c、i_L、i_s 表示与 i_c、v_L 用 v_c、i_L、i_s 表示是一样的. 即在图1所示电路中，将 v_c、i_L、i_s 作为已知量，解出 i_c、v_L 即可，而且利用图2所示电路就更方便了. 在该电路中，将图1所示电路中的电容、电感分别换成电压源 v_c 和电流源 i_L. 这是一个简单的电阻网络，很容易解出 i_c，v_L，由克希霍夫第一定律得

$$i_c(t)+i_L(t)=i_s(t)\Rightarrow i_c(t)=-i_L(t)+i_s(t)$$

由克希霍夫第二定律得

$$v_L(t)+R_2 i_L(t)=v_c(t)\Rightarrow v_L(t)=v_c(t)-R_2 i_L(t)$$

则求得标准形微分方程式

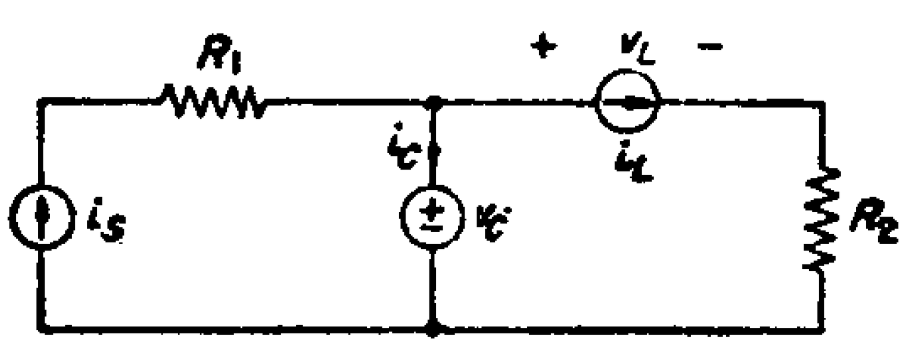


图2　将图1所示电路中的电容换成电压源，电感换成电流源后得到的电路

$$\begin{bmatrix} \dot{v}_c(t) \\ \dot{i}_L(t) \end{bmatrix}=\begin{bmatrix} 0 & -\frac{1}{C} \\ \frac{1}{L} & -\frac{R_2}{L} \end{bmatrix}\begin{bmatrix} v_c(t) \\ i_L(t) \end{bmatrix}+\begin{bmatrix} \frac{1}{C} \\ 0 \end{bmatrix}i_s(t) \tag{4}$$

将此式与(2)式合起来，即可证明$(v_c,\ i_L)$是图1所示电路的状态变量.

归纳以上，得到下列推导电气网络状态方程式的基本方法：

(i) 将电容器上电压、电感中电流选为状态变量；

(ii) 利用这些状态变量和输入量表示电容器的电流和电感上的电压. 如图2所示，这个过程相当于分析电阻网络；

(iii) 在由(ii)得到的结果中，将电容器电流换成 $C\times d$(电容器上电压)$/dt$，将电感电压换成 $L\times d$(电感电流)$/dt$，即可得到标准形微分方程式.

应用上述方法，对于简单电路可以用手工计算求取标准形状态方程式，但是当电路元件的数量一多，整个计算过程就必须利用计算机进行. 为此，必须使(ii)的电阻电路分析

步骤规律化. 这样, 图论知识便是不可缺少的. 而且, 如以后所述, 对于包含有理想变压器、回转器、有源元件等的网络, 其标准形微分方程式的推导过程变得更为复杂, 不能原封不动地使用上述方法. 在这样一般的情况下, 为了得到有规律的统一的方法, 就更必需图论的知识. 为此, 下面将介绍关于图论的基础知识.

图论和克希霍夫定律[30,32-35]

克希霍夫第一、第二定律是网络理论的基础, 该定律将回路中元件上的电压, 电流之间的关系表示成线性方程组的形式. 该线性方程组的形式仅与元件之间的连接关系有关*, 而与元件的种类无关. 因此, 若将所有元件的端子换成节点, 元件换成枝, 由对应的图即可得出该回路的克希霍夫电压方程和电流方程.

元件中电流和电压的方向互相关联, 二者之中只要一个规定, 另一个就自然定了. 现将其基准方向按图 3(a) 确定, 图中枝的方向如图 3(b) 所示. 象这样得到的有向图, 将其枝及节点适当地编以号码, 图中枝和节点的连接关系就可以用矩阵表示, 该矩阵称为关联矩阵 (incidence matrix) $\boldsymbol{A}_a=[a_{ij}]$, 其行对应图的节点, 列对应枝. 各元素的值规定如下:

$$a_{ij}\begin{cases} 1, & \text{枝 } j \text{ 连接在节点 ⓘ 上, 其方向背向节点 ⓘ 时;} \\ 0, & \text{枝 } j \text{ 不连接在节点 ⓘ 上时;} \\ -1, & \text{枝 } j \text{ 连接在节点 ⓘ 上, 其方向指向节点 ⓘ 时.} \end{cases}$$

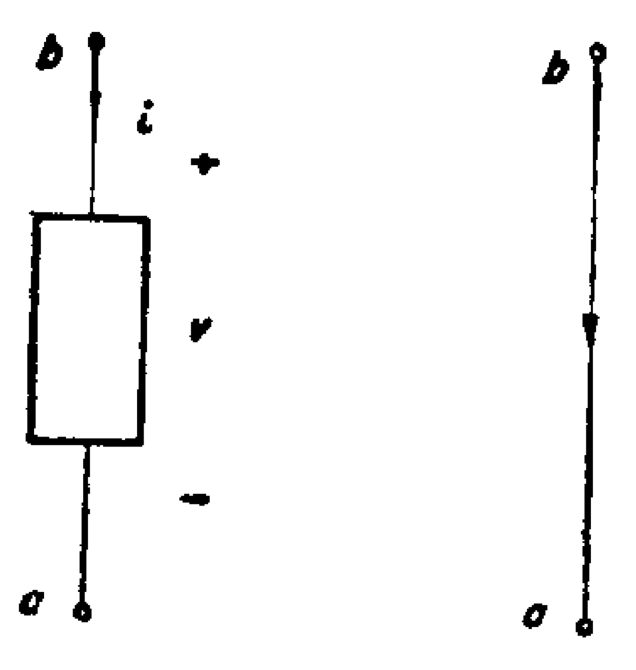


图 3 电压电流方向的相互关系及图中枝的方向

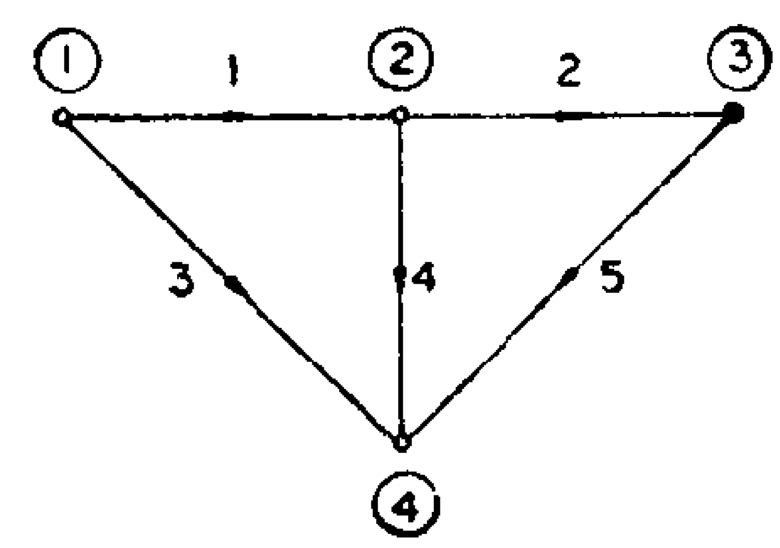


图 4 图 1 所示电路的图

我们来看一个具体例子, 在由图 1 所示电路得到的图中, 将节点和枝适当标以号码, 枝标以方向, 便得到图 4 所示的有向图. 该有向图的关联矩阵是

$$\boldsymbol{A}_a=\begin{array}{c} \overbrace{\begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array}}^{\text{枝}} \\ \left[\begin{array}{rrrrr} 1 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & -1 & -1 & -1 \end{array}\right] \end{array} \left.\begin{array}{c} ① \\ ② \\ ③ \\ ④ \end{array}\right\}\text{节点}$$

在该矩阵的每一列中, 只有一个 1 和一个 -1, 其它元素皆为 0, 这是关联矩阵的共同特点. 如果考虑到一个枝必然由一个节点引出然后进入另一个节点, 这个特点是很容易理

* 仅与网络的几何结构有关. ——译者注

解的. 在这里,若把以图 4 各枝的电流为分量的列向量[1]定义为

$$\boldsymbol{I}_b=[i_1,\ i_2,\ i_3,\ i_4,\ i_5]^T,$$

则对于所有的 t

$$\boldsymbol{A}_a\boldsymbol{I}_b(t)=\boldsymbol{0} \tag{5}$$

成立. 其原因是,将(5)式左边展开后得到的

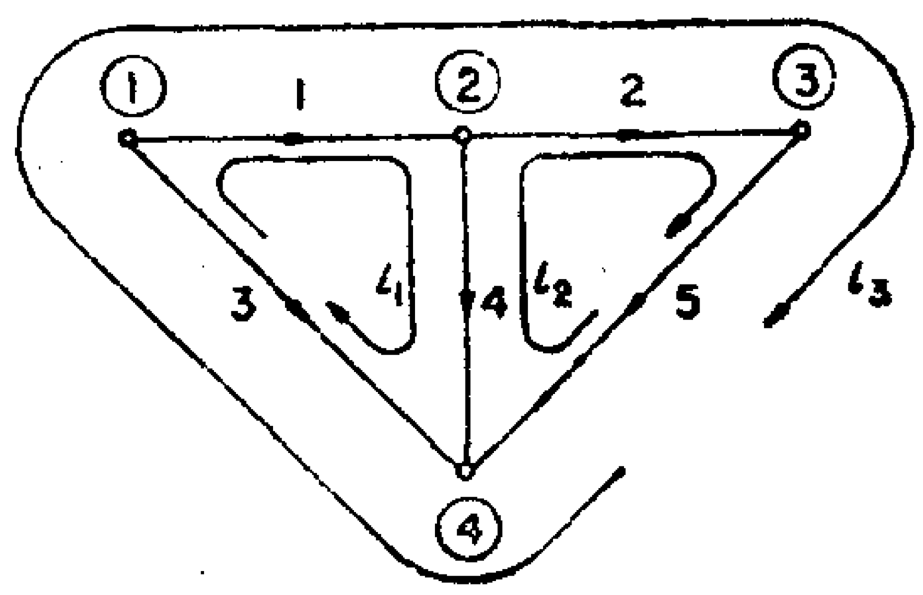


图 5 图 4 所示图的回路

$$\begin{gathered}+{}^*i_1(t)+i_3(t)\\ -i_1(t)+i_2(t)+i_4(t)\\ -i_2(t)+i_5(t)\\ -i_3(t)-i_4(t)-i_5(t)\end{gathered}$$

分别表示由节点 ①②③④ 流出电流的总和,根据克希霍夫第一定律,它们应分别为 0. 根据同样分析,对于一切有向图,其关联矩阵和枝电流均满足关系式(5).

其次,我们来定义回路矩阵 $\boldsymbol{B}_a=[b_{ij}]$. 首先来观察包含在图中所有的回路[2],对于各个回路适当地给以基准方向(每个回路都有两个方向可供选择,可以随便选择其中一个作为基准方向). $\boldsymbol{B}_a$ 的行对应于这样标以方向的回路,列对应于枝,$(i,\ j)$元素 b_{ij} 的值规定如下:

$$b_{ij}=\begin{cases}1, & \text{枝 } j \text{ 在回路 } l_i \text{ 中,其方向与回路 } l_i \text{ 方向一致时;}\\ 0, & \text{枝 } j \text{ 不在回路 } l_i \text{ 中时;}\\ -1, & \text{枝 } j \text{ 在回路 } l_i \text{ 中,但其方向与回路 } l_i \text{ 方向相反.}\end{cases}$$

例如,若将图 1 所示电路的有向图(图 4)中的回路标以图 5 所示方向,则得回路矩阵

$$\boldsymbol{B}_a=\begin{array}{c}\overbrace{\begin{matrix}1 & \ 2 & \ \ 3 & \ \ 4 & 5\end{matrix}}^{\text{枝}}\\ \left[\begin{matrix}1 & 0 & -1 & 1 & 0\\ 0 & 1 & 0 & -1 & 1\\ 1 & 1 & -1 & 0 & 1\end{matrix}\right]\end{array}\begin{array}{l}\\ \left.\begin{matrix}l_1\\ l_2\\ l_3\end{matrix}\right\}\text{回路}\end{array}$$

在这里,若把以该图中各枝的电压为分量的列向量定义为

$$\boldsymbol{V}_b=[v_1,\ v_2,\ v_3,\ v_4,\ v_5]^T,$$

则对于所有的 t

$$\boldsymbol{B}_a\boldsymbol{V}_b(t)=\boldsymbol{0} \tag{6}$$

成立. 其原因是,将式(6)左边展开后得到的

$$\begin{gathered}v_1(t)-v_3(t)+v_4(t)\\ v_2(t)-v_4(t)+v_5(t)\\ v_1(t)+v_2(t)-v_3(t)+v_5(t)\end{gathered}$$

---

[1] 以体 F 的元为元素的 $(n\times 1)$ 矩阵,其乘数仅限于 F 的元,则可以看成 F^n 向量空间的元. 在这个意义上,称它为(属于 F^n 的)列向量(Column Vector). 同样,以体 F 的元为元素的 $(1\times n)$ 矩阵,也可以看成 F^n 的元,称之为行向量(row vector).

* 原文误为 $-i_1(t)+i_3(t)$. ——译者注

[2] 这里所谓的回路是枝和节点连结形成的闭路. 在回路的各节点上都只接有两个枝. 例如 8 字形的闭路不能称为回路.

分别表示沿回路 l_1, l_2, l_3 电压降的总和，根据克希霍夫第二定律，它们应分别为 0. 根据同样的考察可以看到，对于一切有向图，回路矩阵和枝电压均满足关系式(6).

下面将证明，不但由图 4 所示有向图得到的关联矩阵 $\boldsymbol{A}_a$ 和回路矩阵 $\boldsymbol{B}_b$ 满足下列关系

$$\boldsymbol{A}_a\boldsymbol{B}_a^T=\boldsymbol{0},$$

而且对于一切有向图，该关系在其关联矩阵和回路矩阵之间一般均成立.

［定理 1］ 设给定的任意有向图的关联矩阵为 $\boldsymbol{A}_a$, 回路矩阵为 $\boldsymbol{B}_a$, 则

$$\boldsymbol{A}_a\boldsymbol{B}_a^T=\boldsymbol{0} \tag{7}$$

成立.

（证明） 设 $\boldsymbol{A}_a\boldsymbol{B}_a^T$ 的 (i, k) 元素为 c_{ik}, 则

$$c_{ik}=\sum_{j=1}^{n_b} a_{ij}b_{kj}$$

式中 n_b 是所给图的枝数. 若对于 $j=l$, $a_{ij}b_{kj}\neq 0$, 这就意味着枝 l 接在节点 ⓘ 上并且在回路 l_k 中. 因此，根据回路的定义，接在接点 ⓘ 且存在于回路 l_k 中的枝，除了枝 l 以外只还有一个，设该枝为 m. 因为接在接点 ⓘ 且存在于回路 l_k 中的枝，除了枝 l 和 m 以外再没有其它枝了，所以当 $j\neq l$ 且 $j\neq m$ 时，$a_{ij}b_{kj}=0$. 结果

$$c_{ik}=a_{il}b_{kl}+a_{im}b_{km}.$$

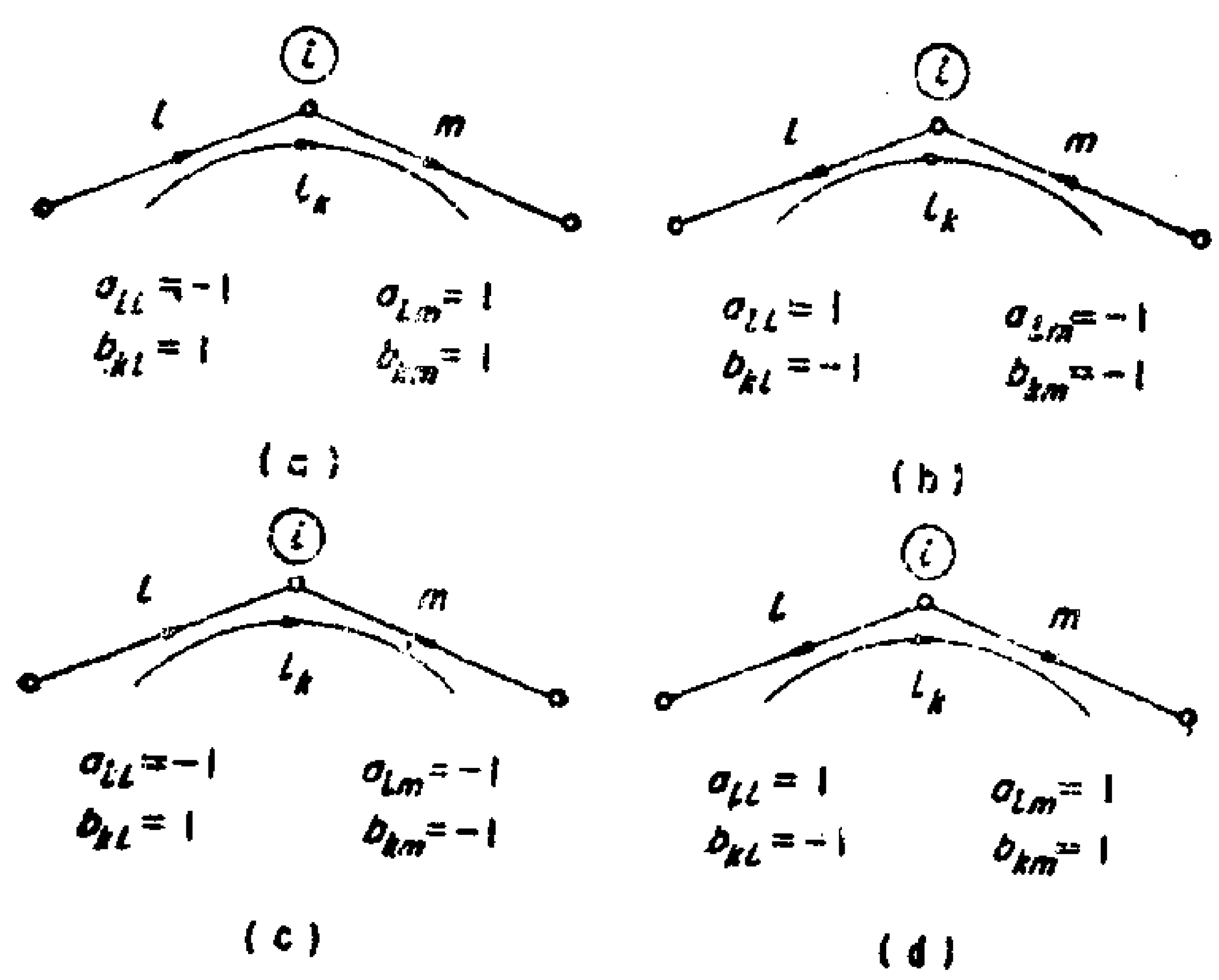


图 6 存在于回路 l_k 中的节点 i 和枝 l, m 的关系

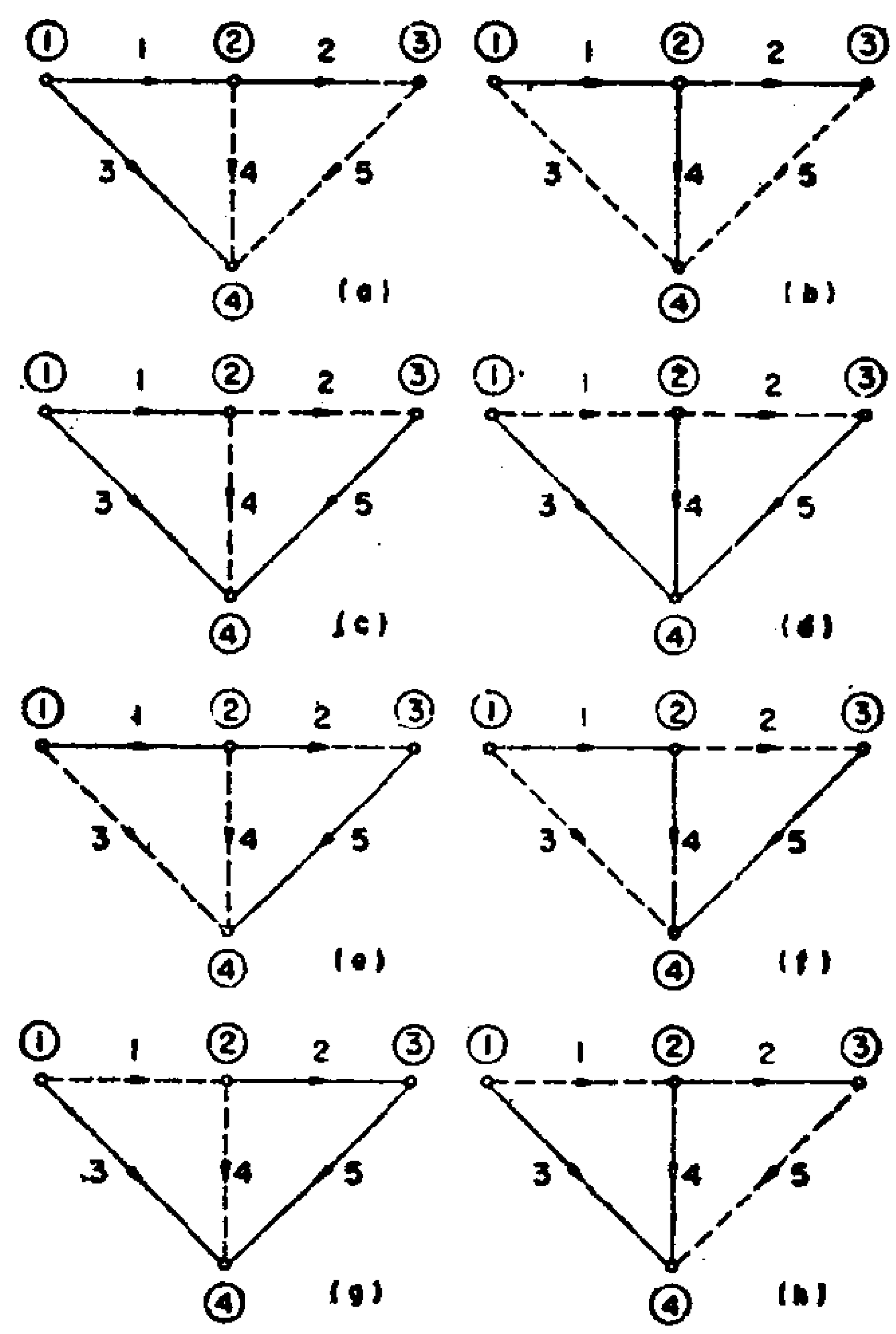


图 7 图 4 图的 8 个树

为了计算 $a_{il}b_{kl}+a_{im}b_{km}$, 只要观察一下图 6 所示的四种情况即可. 显而易见，无论在那一种情况下，均存在 $a_{il}b_{kl}+a_{im}b_{km}=0$. 以上表明，对于任意的 i, k, 均存在 $c_{ik}=0$, 这就说明(7)式在一般情况下均成立. （证明完毕）

下面我们来定义树(tree)和树余(cotree). 为了简单起见，设所讨论的图为连结图，即在任意两个节点之间至少有一个连结通道(连结枝).

［定义 1］ 连结给定图中全部节点的最少数量枝的集合称为树，在所给图形中除掉树后余下枝的集合称为树余.

由该定义可见，树余决定于所选的树. 对于给定图，一般树不是唯一的，因而树余也不是唯一确定的. 图 4 所示图有 8 个树，示于图 7. 设图的节点数为 n_n, 枝数为 n_b, 则树

具有(n_n-1)个枝，树余具有(n_b-n_n+1)个枝.

其次，对于连结性图我们来定义割集(cut set).

[定义 2] 在给定的连结性图中，将某些枝取掉(切断)后，原图形便分成两个连结的部分图，即使被取掉(切断)的一个枝留下，图仍保持其连结性，这种被取掉(切断)最小数目枝的集合，称为割集.

图 4 所示图有 6 个割集，现用枝号码的集合表示如下：

$$c_1:\{1,\ 3\} \qquad c_2:\{1,\ 2,\ 4\}$$
$$c_3:\{2,\ 5\} \qquad c_4:\{3,\ 4,\ 5\}$$
$$c_5:\{1,\ 4,\ 5\} \qquad c_6:\{2,\ 3,\ 4\}$$

(注意) 例如取掉枝 1, 3 后，整个图被分成两块，一块仅包含节点①，另一块包含节点②③④及枝 2, 4, 5. 因此，枝 1, 3 是割集. 要注意，一个节点(在该例中是节点①)也可以看成是原图的部分连结性图.

象回路可以用回路矩阵表示一样，我们也要用矩阵表示割集. 为此，我们来观察一下图中所有的割集，并按下列方法给这些割集以适当的基准方向：每个割集均将原图分成两个部分图，设其为 G_1 和 G_2. 我们规定从 G_1 到 G_2 (或从 G_2 到 G_1) 的方向为各割集的方向. 割集矩阵 $\boldsymbol{Q}_a=[q_{ij}]$是行对应于按上述方法规定了方向的割集，列对应于枝的矩阵，$(i,\ j)$元素的值 q_{ij} 规定如下：

$$q_{ij}=\begin{cases}1, & \text{枝 } j \text{ 包含在割集 } c_i \text{ 中，其方向与割集 } c_i \text{ 的方向一致时}^*;\\ 0, & \text{枝 } j \text{ 不包含在割集 } c_i \text{ 中时};\\ -1, & \text{枝 } j \text{ 包含在割集 } c_i \text{ 中，但其方向与 } c_i \text{ 的方向相反时}.\end{cases}$$

将图 4 所示有向图的割集适当地标以方向，得

$$\boldsymbol{Q}_a=\begin{array}{c}\overbrace{\begin{array}{ccccc}1 & 2 & 3 & 4 & 5\end{array}}^{\text{枝}}\\ \left[\begin{array}{rrrrr}1 & 0 & 1 & 0 & 0\\ -1 & 1 & 0 & 1 & 0\\ 0 & -1 & 0 & 0 & 1\\ 0 & 0 & -1 & -1 & -1\\ 1 & 0 & 0 & -1 & -1\\ 0 & -1 & -1 & -1 & 0\end{array}\right]\end{array}\left.\begin{array}{c}c_1\\ c_2\\ c_3\\ c_4\\ c_5\\ c_6\end{array}\right\}\text{割集}$$

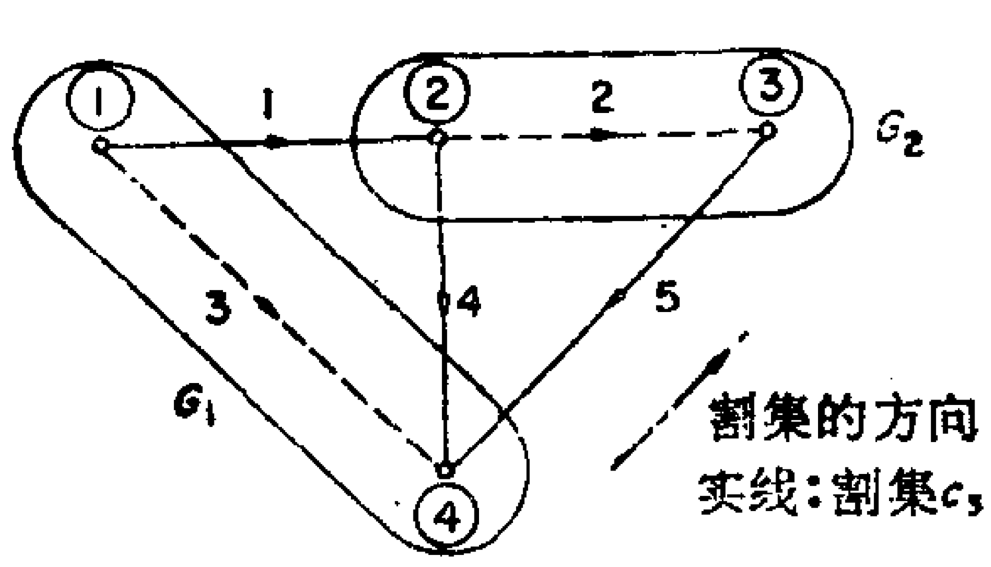


图8 图 4 所示图的割集 c_5

例如，其中割集 c_5 的方向取为 G_1 到 G_5，如图 8 所示.

将 $\boldsymbol{Q}_a$ 和前面讲过的关联矩阵 $\boldsymbol{A}_a$ 进行比较，我们发现，$\boldsymbol{Q}_a$ 的每一行都是 $\boldsymbol{A}_a$ 的某几行之和，例如，$\boldsymbol{Q}_a$ 的第 5 行是 $\boldsymbol{A}_a$ 的第 1 行和第 4 行之和. 下面将证明，这个关系在 $\boldsymbol{Q}_a$ 和 $\boldsymbol{A}_a$ 之间一般均成立.

[定理 2] 设给定的任意(连结的)有向图的关

* 每一个割集 c_i 均将给定图的节点集合 N 分成两个部分 N_1 和 N_2. 上面已规定 c_i 的方向为由 G_1 到 G_2(N_1 属于 G_1，N_2 属于 G_2)，当枝 $\boldsymbol{b}_i$ 的始点属于 N_1，终点属于 N_2 时. 我们称为枝 $\boldsymbol{b}_i$ 和割集 c_i 的方向一致，当枝 $\boldsymbol{b}_i$ 的始点属于 N_2，终点属于 N_1 时，称为 $\boldsymbol{b}_i$ 和 c_i 方向相反. ——译者注

联矩阵为 $\boldsymbol{A}_a$，割集矩阵为 $\boldsymbol{Q}_a$，则 $\boldsymbol{Q}_a$ 的任意一行均可以用 $\boldsymbol{A}_a$ 的某几个适当行之和表示. 因此，利用适当的矩阵 $\boldsymbol{W}_a$ 可以表示成

$$\boldsymbol{Q}_a=\boldsymbol{W}_a\boldsymbol{A}_a \tag{8}$$

（证明） 假定割集 c_i 将给定的连结性图分成两个连结性的部分图 G_1G_2，我们规定该割集 c_i 的方向为由 G_1 指向 G_2，现在来考察 $\boldsymbol{A}_a$ 中由位于 G_1 中节点对应的行所组成的部分矩阵. 在该部分矩阵中，G_1 所包含的枝对应的列上，1 和 -1 刚好各一个；组成割集的枝所对应的列上，只有一个 1 或 -1；G_2 包含的枝所对应的列上，所有的元素均为 0. 因此，在该部分矩阵行的总和形成的行向量中，除了与组成割集的枝对应的元素以外，其它均为 0. 而且，因为规定割集的方向为由 G_1 指向 G_2，组成割集的枝所对应的元素，当枝的方向为由 G_1 指向 G_2 时为 1，由 G_2 指向 G_1 时为 -1，所以该行向量与 $\boldsymbol{Q}_a$ 的割集 c_i 对应的行一致.

对于图 4 所示有向图，有

$$\begin{array}{c}
\overbrace{\begin{array}{ccccc}1 & 2 & 3 & 4 & 5\end{array}}^{\text{枝}} \\
\underbrace{\left[\begin{array}{ccccc}
1 & 0 & 1 & 0 & 0 \\
-1 & 1 & 0 & 1 & 0 \\
0 & -1 & 0 & 0 & 1 \\
0 & 0 & -1 & -1 & -1 \\
1 & 0 & 0 & -1 & -1 \\
0 & -1 & -1 & -1 & 0
\end{array}\right]}_{\boldsymbol{Q}_a}
\end{array}
\begin{array}{l}
\\
\left.\begin{array}{l} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{array}\right\}\text{割集}
\end{array}$$

$$=\underbrace{\left[\begin{array}{cccc}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
1 & 0 & 0 & 1 \\
0 & 0 & 1 & 1
\end{array}\right]}_{\boldsymbol{W}_a}
\begin{array}{c}
\overbrace{\begin{array}{ccccc}1 & 2 & 3 & 4 & 5\end{array}}^{\text{枝}} \\
\underbrace{\left[\begin{array}{ccccc}
1 & 0 & 1 & 0 & 0 \\
-1 & 1 & 0 & 1 & 0 \\
0 & -1 & 0 & 0 & 1 \\
0 & 0 & -1 & -1 & -1
\end{array}\right]}_{\boldsymbol{A}_a}
\end{array}
\begin{array}{l}
\\
\left.\begin{array}{l} ① \\ ② \\ ③ \\ ④ \end{array}\right\}\text{节点}
\end{array}$$

该定理很重要，和(5)，(7)式结合起来，得

$$\boldsymbol{Q}_a\boldsymbol{I}_b(t)=\mathbf{0}\quad \forall t \tag{9}$$

$$\boldsymbol{Q}_a\boldsymbol{B}_a^T(t)=\mathbf{0} \tag{10}$$

(9)式表明，对于任何割集电流总和在所有的时间 t 均为 0. 换句话说，它表明从被闭合曲线所包围的回路中流出电流的总和，在所有的时间 t 均为 0. (9)式称为一般情况下的克希霍夫第一定律，(5)式是它的特殊情况.

其次，我们来定义在回路和割集的概念中特别重要的基本回路和基本割集，并证明其之间的重要关系.

［定义 3］ 根据树的定义，在树上加上其树余的任何一个枝，都必然形成一个回路，这种回路称为基本回路(fundamental loop). 由回路矩阵中仅取出行与基本回路对应的

部分矩阵,称为基本回路矩阵.

按照从树枝到树余枝的顺序对枝进行编号,若将存在于各基本回路中树余枝的方向取为该基本回路的方向,则基本回路矩阵可写成

$$\boldsymbol{B}_f=[\overbrace{\boldsymbol{B}_1,}^{n_n-1}\ \overbrace{\ \boldsymbol{I}\ }^{n_b-n_n+1}]\}n_b-n_n+1 \tag{11}$$

式中 $\boldsymbol{I}$ 是单位矩阵. 与枝的编号相对应,设以树的枝电压为分量的电压向量为 $\boldsymbol{V}_t$, 以树余的枝电压为分量的电压向量为 $\boldsymbol{V}_c$, 则整个电压向量可表示成

$$\boldsymbol{V}_b=\begin{bmatrix}\boldsymbol{V}_t\\ \boldsymbol{V}_c\end{bmatrix}\begin{matrix}\}n_n-1\\ \}n_b-n_n+1\end{matrix}$$

因 $\boldsymbol{B}_f$ 是回路矩阵 $\boldsymbol{B}_a$ 的部分矩阵,则由(6)式可知,对于一切 t,

$$\boldsymbol{B}_f\boldsymbol{V}_b(t)=\boldsymbol{0}$$

成立. 利用(11)式,根据 B-I-1 中表 1, 将乘式展开,得

$$\boldsymbol{V}_c(t)=-\boldsymbol{B}_1\boldsymbol{V}_t(t) \tag{12}$$

由该式可见,树余的枝电压可以用树的枝电压表示.

此外,由树的定义显而易见,树的任意一个枝适当地加上树余的枝,就能变成割集. 注意到这个事实,基本割集就可定义如下:

[定义 4] 所谓基本割集*(fundamental cut set),是由一个树枝和尽可能少的树余的枝组成的割集. 在割集矩阵中仅取出与基本割集对应行的部分矩阵,该部分矩阵叫做基本割集矩阵.

若各割集的方向作为包含在割集中树枝的方向, 利用在讲述基本回路矩阵时枝的编号,则基本割集矩阵 $\boldsymbol{Q}_f$ 可以写成下列形式

$$\boldsymbol{Q}_f=[\overbrace{\boldsymbol{I},}^{n_n-1}\ \overbrace{\ \boldsymbol{Q}_1\ }^{n_b-n_n+1}]\}n_n-1 \tag{13}$$

因为 $\boldsymbol{Q}_f$ 是割集矩阵 $\boldsymbol{Q}_a$ 的部分矩阵,根据(9)式,对于一切 t. 则

$$\boldsymbol{Q}_f\boldsymbol{I}_b(t)=\boldsymbol{0}$$

成立. 设 $\boldsymbol{I}_t$ 是以树的枝电流为分量的电流向量, $\boldsymbol{I}_c$ 为以树余的枝电流为分量的电流向量. 因为全电流向量 $\boldsymbol{I}_b$ 可以写成下列形式

$$\boldsymbol{I}_b=\begin{bmatrix}\boldsymbol{I}_t\\ \boldsymbol{I}_c\end{bmatrix}\begin{matrix}\}n_n-1\\ \}n_b-n_n+1\end{matrix}$$

根据 B-I-1 表 1, 对于一切 t, 则下式成立:

$$\boldsymbol{Q}_f\boldsymbol{I}_b(t)=\boldsymbol{I}_t(t)+\boldsymbol{Q}_1\boldsymbol{I}_c(t)=\boldsymbol{0}$$

将该式变化一下,便得到基本割集方程式

$$\boldsymbol{I}_t(t)=-\boldsymbol{Q}_1\boldsymbol{I}_c(t) \tag{14}$$

由这个关系可见,树的枝电流可以用树余的枝电流表示. 那么,由基本回路矩阵的形式可见, (12) 式是 (n_b-n_n+1) 个独立方程式; 由基本割集矩阵的形式可见, (14) 式是 (n_n-1) 个独立方程式. 这些方程式和各枝的端子电压-电流特性方程式 (n_b 个) 加到一起, 总共得到 $2n_b$ 个独立方程式, 而这个数目等于所给网络未知数(各枝的端子电压和电流)的总数, 因此, 这些方程可以解出. 这个事实说明, 可以代替克希霍夫第二定律的式

* 它是相对于某个树定义的. ——译者注

(6)，和普遍的克希霍夫第一定律的式(9)，使用方程数目比它们还少的基本回路方程(12)和基本割集合方程(14)。

虽然割集矩阵 $\boldsymbol{Q}_a$ 和回路矩阵 $\boldsymbol{B}_a$ 之间的关系式(10)对于基本割集矩阵 $\boldsymbol{Q}_f$ 和基本回路矩阵 $\boldsymbol{B}_f$ 也成立，但是利用它可以推出 $\boldsymbol{Q}_f$ 和 $\boldsymbol{B}_f$ 之间的关系，即若利用(11)和(13)式，则

$$\boldsymbol{Q}_f\boldsymbol{B}_f^T=[\boldsymbol{I},\ \boldsymbol{Q}_1]\begin{bmatrix}\boldsymbol{B}_1^T\\ \boldsymbol{I}\end{bmatrix}$$

参照 B-I-1 中表 1，得

$$\boldsymbol{Q}_f\boldsymbol{B}_f^T=\boldsymbol{B}_1^T+\boldsymbol{Q}_1=\mathbf{0}.$$

因此，若令

$$\boldsymbol{Q}_1=-\boldsymbol{B}_1^T\triangleq\boldsymbol{F} \tag{15}$$

则可表示成

$$\boldsymbol{B}_f=[-\boldsymbol{F}^T,\ \boldsymbol{I}] \tag{16a}$$

$$\boldsymbol{Q}_f=[\boldsymbol{I},\ \boldsymbol{F}]. \tag{16b}$$

由此可见，$\boldsymbol{B}_f$ 和 $\boldsymbol{Q}_f$ 之中只要求出一个，立即可得到另外一个。

(问题 1) 在所给图上，任意回路都可以用几个基本回路的组合表示，用矩阵的术语来说，即回路矩阵 $\boldsymbol{B}_a$ 的任意行都可以用基本回路矩阵 $\boldsymbol{B}_f$ 的行线性表出，试证明之。

(问题 2) 在所给图上，任意割集都可以用几个基本割集的组合表示，即割集矩阵 $\boldsymbol{Q}_a$ 的任意行都可以用基本割集矩阵 $\boldsymbol{Q}_f$ 的行线性表出，试证明之。

最后，再来谈一下关联矩阵 $\boldsymbol{A}_a$ 的秩，说明由 $\boldsymbol{A}_a$ 求写成(13)式形式的基本割集矩阵 $\boldsymbol{Q}_f$ 的方法。若 $\boldsymbol{Q}_f$ 已求出，与其最前面的 (n_n-1) 列相对应的枝构成树，如(16)式所示，因为基本回路矩阵 $\boldsymbol{B}_f$ 立即可求出，所以该方法很重要。为了简单起见，设所给图是连结性图。注意，在连结性图上，$n_n-1\leqslant n_b$。

如前所述，因为在 $n_n\times n_b$ 矩阵 $\boldsymbol{A}_a$ 的各列中，都各有一个为 1 和一个为 -1 的元素，其它元素全为零，所以 $\boldsymbol{A}_a$ 的各行之总和是零行向量。这意味着，$\boldsymbol{A}_a$ 的各行不是线性独立的。因此，根据 B-I-3，说明 $\boldsymbol{A}_a$ 的秩在 n_n-1 以下。

现在，对图中的枝和节点适当加以编号，设枝 1 到枝 (n_n-1) 组成树，即设与关联矩阵 $\boldsymbol{A}_a$ 的第 1 列～第 (n_n-1) 列相对应的枝组成树。这时如(8)式所示，利用适当的 $(n_n-1)\times n_n$ 矩阵 $\boldsymbol{W}_1$，可将 $(n_n-1)\times n_b$ 矩阵 $\boldsymbol{Q}_f$ 表示成

$$\boldsymbol{Q}_f=\boldsymbol{W}_1\boldsymbol{A}_a.$$

枝的号码与 $\boldsymbol{A}_a$ 列的排列相对应。因为枝编号的变化是 $\boldsymbol{A}_a$ 列的顺序号码的变化，所以对枝号任意编的图，利用适当的顺列矩阵 $\boldsymbol{Z}_2=(n_b\times n_b)$，可以表示成

$$\boldsymbol{Q}_f=\boldsymbol{W}_1\boldsymbol{A}_a\boldsymbol{Z}_2 \tag{17}$$

由(13)式 $\boldsymbol{Q}_f$ 的结构可见，$\boldsymbol{Q}_f$ 的秩是 n_n-1，由 B-I-3 中式(23)得知，$\boldsymbol{A}_a$ 的秩在 n_n-1 以上，因此，连结性图的关联矩阵的秩是 n_n-1。

现在，我们来看由 $\boldsymbol{A}_a$ 中取掉任意一行的 $(n_n-1)\times n_b$ 矩阵 $\boldsymbol{A}$，如上所述，因为这个取出的行与 $\boldsymbol{A}$ 的所有各行之和是零行向量，所以该行可用 $\boldsymbol{A}$ 的行线性表出。而且，利用适当的 $n_n\times(n_n-1)$ 矩阵 $\boldsymbol{W}_2$，可将 $\boldsymbol{A}_a$ 表示成

$$\boldsymbol{A}_a=\boldsymbol{W}_2\boldsymbol{A} \tag{18}$$

因为 $\boldsymbol{A}_a$ 的秩是 n_n-1，另外由 B-I-3 中(23)式，$\boldsymbol{A}$ 的秩不能少于 n_n-1，又因 $\boldsymbol{A}$ 是 $\boldsymbol{A}_a$ 的部分矩阵，其秩不能大于 n_n-1，因此，$\boldsymbol{A}$ 的秩是 n_n-1。

将(17)和(18)式结合起来，$\boldsymbol{Q}_f$ 可以表示成

$$\boldsymbol{Q}_f = \boldsymbol{W}_1\boldsymbol{W}_2\boldsymbol{A}\boldsymbol{Z}_2 \triangleq \boldsymbol{Z}_1\boldsymbol{A}\boldsymbol{Z}_2 \tag{19}$$

式中 $\boldsymbol{Z}_1 \triangleq \boldsymbol{W}_1\boldsymbol{W}_2 = (n_n-1)\times(n_n-1)$. 因为 $\boldsymbol{Q}_f$ 的秩是 n_n-1，则即便在这里也可以使用 B-I-3 中(23)式，可见 $\boldsymbol{Z}_1$ 必须是正则的. 根据 B-I-2 中系 4，$\boldsymbol{Z}_1$ 可以表示成有限个基本行变换矩阵之积. 因此，(19)式表明 $\boldsymbol{Q}_f$ 是由 $\boldsymbol{A}$ 经过基本行变换和列交换后得到的. 即利用基本行变换和列交换，将 $\boldsymbol{A}$ 的第 1 列～第(n_n-1)列作成单位矩阵便得到 $\boldsymbol{Q}_f$. 以上结论稍加整理便得到下列定理.

[定理 3] 在连结性图关联矩阵 $\boldsymbol{A}_a$ 中取掉任意一行，剩下的$(n_n-1)\times n_b$ 矩阵 $\boldsymbol{A}$ 的秩为(n_n-1)，对 $\boldsymbol{A}$ 施行适当的基本行变换和列交换，便可得到基本割集矩阵 $\boldsymbol{Q}_f$. 具体步骤如下：

(i) 将 $\boldsymbol{A}$ 变换成行标准形(参看 B-I-2)；

(ii) 在所得到的行标准形上，适当进行列交换，使第 1 列～第 (n_n-1) 列构成单位矩阵 $\boldsymbol{I}_{n_n-1}$.

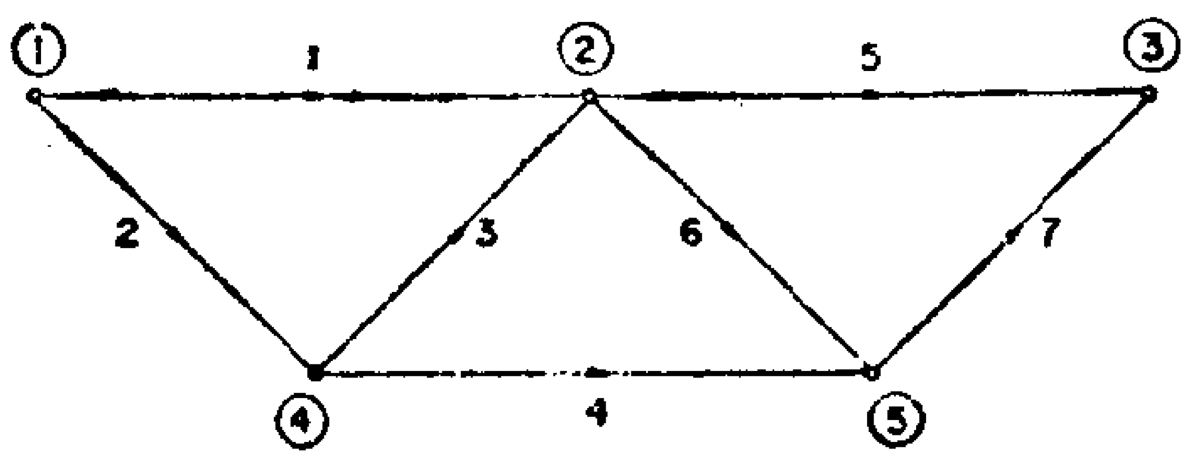


图9 例1

为了区别矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{A}_a$，矩阵 $\boldsymbol{A}$ 通常叫做既约关联矩阵(reduced incidence matrix). 下面举个实际例子，看一下由 $\boldsymbol{A}_a$ 推导 $\boldsymbol{Q}_f$ 的过程.

(例 1) 图 9 所示图的关联矩阵为

$$\boldsymbol{A}_a = \begin{array}{c} \overbrace{\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array}}^{\text{枝}} \\ \left[\begin{array}{ccccccc} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & -1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & -1 \\ 0 & -1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & -1 & 1 \end{array}\right] \end{array} \begin{array}{c} \\ \left.\begin{array}{c} ① \\ ② \\ ③ \\ ④ \\ ⑤ \end{array}\right\} \text{节点}, \end{array}$$

消去 $\boldsymbol{A}_a$ 中任意一行，例如第 5 行，设所得到的既约关联矩阵为 $\boldsymbol{A}$，

$$\boldsymbol{A} = \begin{array}{c} \overbrace{\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array}}^{\text{枝}} \\ \left[\begin{array}{ccccccc} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & -1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & -1 \\ 0 & -1 & 1 & 1 & 0 & 0 & 0 \end{array}\right] \end{array},$$

其次，将 $\boldsymbol{A}$ 变成标准形(参看 B-I-2 中的步骤). 把第 1 行加到第 2 行，得

$$\begin{array}{c} \overbrace{\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array}}^{\text{枝}} \\ \left[\begin{array}{ccccccc} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & -1 \\ 0 & -1 & 1 & 1 & 0 & 0 & 0 \end{array}\right] \end{array},$$

从第1行减去第2行；再把第2行加到第4行，得

$$\begin{array}{c} \overbrace{\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array}}^{\text{枝}} \\ \begin{bmatrix} 1 & 0 & 1 & 0 & -1 & -1 & 0 \\ 0 & 1 & -1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix} \end{array},$$

第3行乘以(−1)后加到第1行；再分别由第2行及第4行减去第3行乘以(−1)，得

$$\begin{array}{c} \overbrace{\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array}}^{\text{枝}} \\ \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & -1 & 1 \\ 0 & 1 & -1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & -1 \end{bmatrix} \end{array},$$

将第3行和第4行交换，得

$$\begin{array}{c} \overbrace{\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array}}^{\text{枝}} \\ \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & -1 & 1 \\ 0 & 1 & -1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix} \end{array},$$

这里便得到了行标准形．其次，利用列交换使矩阵左端作成4×4阶单位矩阵即可，即将第3列移到右端，第4、5、6、7列顺序左移一列，便得到Q_f

$$Q_f = \begin{array}{c} \overbrace{\begin{array}{ccccccc} 1 & 2 & 4 & 5 & 6 & 7 & 3 \end{array}}^{\text{枝}} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & -1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} \end{array}.$$

由此可见，枝1、2、4、5组成一个树(参看图10)．一个图一般有几个树，它们都可以用上述变换关联矩阵A_a列的顺序的方法得到．

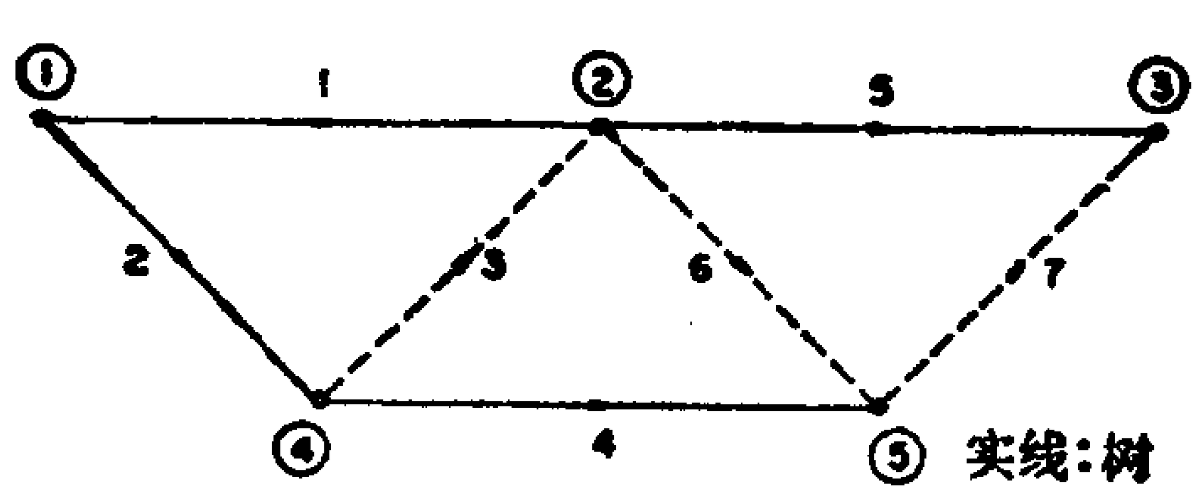


图10　例1图的树(一个例子)

因该方法仅由基本行变换和列交换操作组成，复杂图的树以及基本割集矩阵也比较容易找到，则是适合于计算机的方法.

非时变线性 *RLC* 网络的状态方程式[32,36]

在这里，首先对于简单的 RLC 网络基本状态方程式的推导再进行较图 1 更为一般性的讨论. 其次，来讨论线性非时变 RLC 网络状态方程式更为系统的推导方法.

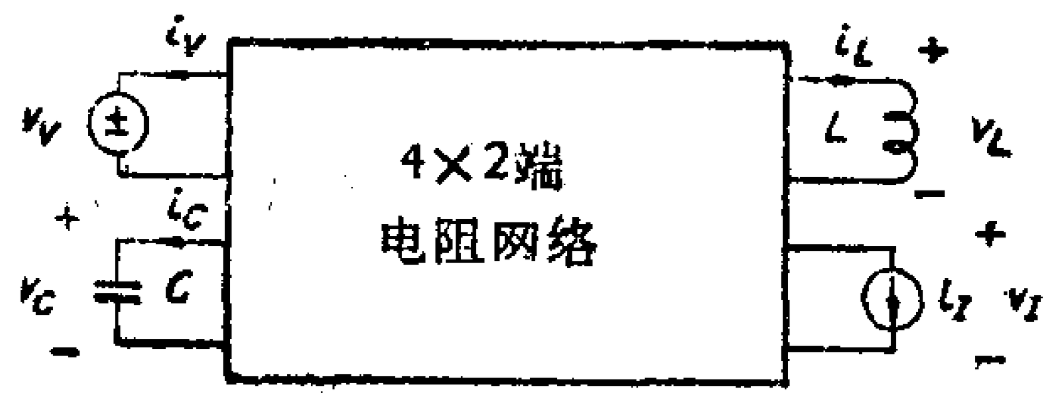


图 11 具有电容、电感、电压源、电流源各 1 个的 RLC 网络

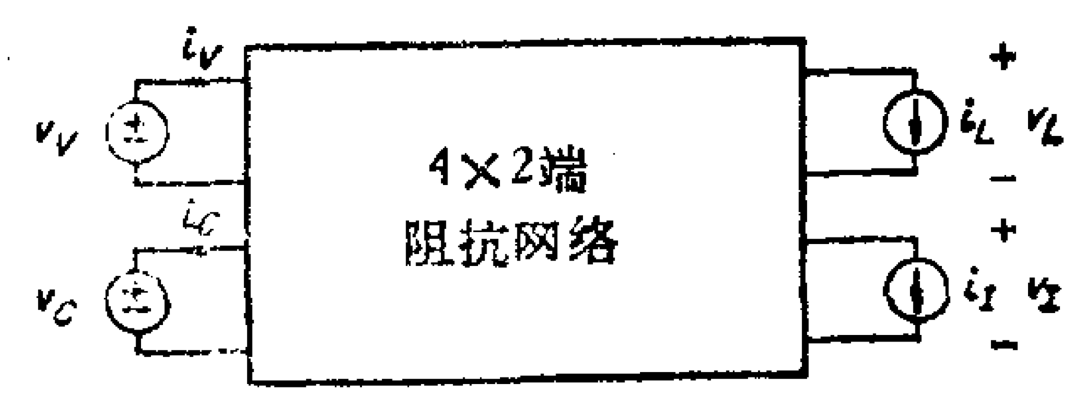


图 12 将图 11 RLC 网络的电容换成电压源，电感换成电流源后得到的阻抗网络

现在，我们先来讨论由电压源、电流源、电容、电感各 1 个，电阻若干个组成的 RLC 网络. 这里假定电压源和电容不并联，电流源和电感不串联. 在电路中将电压源、电流源、电容、电感提出来，可以表示成图 11 的形式. 将电容上电压 v_c 和电感中电流 i_L 看成形式上独立的变量，如图 12 所示，电容可以换成电压源，电感可以换成电流源(参看图 12). 这时，除掉电压源、电流源后的 4×2 端阻抗网络*，其外特性可以用下列混合参数表示[32]**：

$$\begin{bmatrix} i_C(t) \\ v_L(t) \\ i_V(t) \\ v_I(t) \end{bmatrix} = \begin{bmatrix} h_{CC} & h_{CL} & h_{CV} & h_{CI} \\ h_{LC} & h_{LL} & h_{LV} & h_{LI} \\ h_{VC} & h_{VL} & h_{VV} & h_{VI} \\ h_{IC} & h_{IL} & h_{IV} & h_{II} \end{bmatrix} \begin{bmatrix} v_C(t) \\ i_L(t) \\ v_V(t) \\ i_I(t) \end{bmatrix} \tag{20}$$

从第 1 行和第 2 行得

$$\begin{bmatrix} i_C(t) \\ v_L(t) \end{bmatrix} = \begin{bmatrix} h_{CC} & h_{CL} \\ h_{LC} & h_{LL} \end{bmatrix} \begin{bmatrix} v_C(t) \\ i_L(t) \end{bmatrix} + \begin{bmatrix} h_{CV} & h_{CI} \\ h_{LV} & h_{LI} \end{bmatrix} \begin{bmatrix} v_V(t) \\ i_I(t) \end{bmatrix},$$

将电容和电感的电压——电流特性式

$$\begin{bmatrix} i_C(t) \\ v_L(t) \end{bmatrix} = \begin{bmatrix} C & 0 \\ 0 & L \end{bmatrix} \begin{bmatrix} \dot{v}_C(t) \\ \dot{i}_L(t) \end{bmatrix}$$

代入上式，得

$$\begin{bmatrix} C & 0 \\ 0 & L \end{bmatrix} \begin{bmatrix} \dot{v}_C(t) \\ \dot{i}_L(t) \end{bmatrix} = \begin{bmatrix} h_{CC} & h_{CL} \\ h_{LC} & h_{LL} \end{bmatrix} \begin{bmatrix} v_C(t) \\ i_L(t) \end{bmatrix} + \begin{bmatrix} h_{CV} & h_{CI} \\ h_{LV} & h_{LI} \end{bmatrix} \begin{bmatrix} v_V(t) \\ i_I(t) \end{bmatrix}$$

于是，得到 $\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{u}$ 形的标准形微分方程式

$$\begin{bmatrix} \dot{v}_C(t) \\ \dot{i}_L(t) \end{bmatrix} = \begin{bmatrix} \dfrac{h_{CC}}{C} & \dfrac{h_{CL}}{C} \\ \dfrac{h_{LC}}{L} & \dfrac{h_{LL}}{L} \end{bmatrix} \begin{bmatrix} v_C(t) \\ i_L(t) \end{bmatrix} + \begin{bmatrix} \dfrac{h_{CV}}{C} & \dfrac{h_{CI}}{C} \\ \dfrac{h_{LV}}{L} & \dfrac{h_{LI}}{L} \end{bmatrix} \begin{bmatrix} v_V(t) \\ i_I(t) \end{bmatrix} \tag{21}$$

因为图 12 所示 4×2 端阻抗网络内部的电压、电流可以唯一地表示成外部电压源的

* 4×2 端阻抗网络即四端子对电阻网络. ——译者注

** 由下面利用图论推导状态方程式的过程也很容易理解这一点. ——译者注

电压和电流源的电流的线性阻合[32]*. 由(20)式可见，电压源的电流和电流源的电压也可以同样表示，所以在所研究的电路中无论把那一个变量(电压、电流)作为输出量，都可以表示成 $\boldsymbol{y}=\boldsymbol{Cx}+\boldsymbol{Du}$ 的形式. 因此，$\boldsymbol{x}=[v_C,\ i_L]^T$ 满足状态变量的条件 (P.2)，(P.3)，(21)式是状态方程式. 由上述可见，RLC 网络状态方程式的推导过程可以归纳如下：从给定网络中将电压源、电流源、电容、电感提出来，然后求剩下的 $N\times 2$ 端阻抗网络的混合参数特性方程式.

应用前面讲过的图论知识，实际上也求出线性非时变 RLC 网络的状态方程式. 这里在讨论 RLC 网络时认为 R、L、C 均取正值，且下列假定成立：

A(i) 没有只有电压源的回路；

A(ii) 没有只有电流源的割集；

A(iii) 没有只有电压源和电容的回路；

A(iv) 没有只有电容的回路；

A(v) 没有只有电流源和电感的割集；

A(vi) 没有只有电感的割集.

在所讨论的网络中，不存在只具有电压源的回路和只具有电流源的割集. 否则，由于包含在这种回路(割集)中的电压源(电流源)的独立性，会产生物理学上的矛盾，亦即不一定满足克希霍夫定律. 既使偶尔满足克希霍夫定律，网络的性能也不会改变. 这是因为，除掉几个电压源、电流源，这样的回路和割集不存在. 由于电容电压、电感电流都可选作状态变量，则在假定 A(iii)—(vi)之下，就可以利用上述基本原则，对于这些假定不成立的场合，以后再讨论.

根据上述基本原则，电容电压和电感电流在形式上看成独立变量，如(12)和(14)式所示. 因为树余的枝电压和树的枝电流是因变量，故为了把电容电压和电感电流作为独立变量，必须这样选择树，使电容枝包含在树中，电感枝包含在树余中. 由于电压源的电压和电流源的电流也是独立变量，所以根据同样理由，在选择树时必须使电压源枝包含在树中，电流源枝包含在树余中. 根据假定 A(i)—A(vi)，在这里所讨论的 RLC 网络中，通常总可以按上述原则选择树，因此给出下列定义.

[定义 5] 包含了所有的电压源枝和电容枝而完全不包含电流源枝和电感枝的树，叫做常态树(proper tree，写成 P 树).

表1 枝的分类

枝的种类	电压向量	电流向量	枝的数量	
电压源枝	$\boldsymbol{V}_V$	$\boldsymbol{I}_V$	n_V	树
电容枝	$\boldsymbol{V}_C$	$\boldsymbol{I}_C$	n_C	
包含在树中的电阻枝	$\boldsymbol{V}_G$	$\boldsymbol{I}_G$	n_G	
包含在树余中的电阻枝	$\boldsymbol{V}_R$	$\boldsymbol{I}_R$	n_R	树余
电感枝	$\boldsymbol{V}_L$	$\boldsymbol{I}_L$	n_L	
电流源枝	$\boldsymbol{V}_I$	$\boldsymbol{I}_I$	n_I	

* 见上页注**.

实际上，为了产生 P 树，在图中按着电压源枝、电容枝、电阻枝、电感枝、电流源枝的顺序，和既约关联矩阵 $\boldsymbol{A}$ 的第 1 列、第 2 列、……相对应地确定 $\boldsymbol{A}$，这时可利用定理 3 的步骤.

以这样的树为基础，按电压源枝、电容枝、包含在树中的电阻枝、包含在树余中的电阻枝、电感枝、电流源枝的顺序对枝进行编号. 表 1 所示为以各枝的电压、电流为分量的向量. 整个电压向量 $\boldsymbol{V}_b$ 和电流向量 $\boldsymbol{I}_b$ 可表示如下:

$$\boldsymbol{V}_b=\begin{bmatrix}\boldsymbol{V}_V\\\boldsymbol{V}_C\\\boldsymbol{V}_G\\\boldsymbol{V}_R\\\boldsymbol{V}_L\\\boldsymbol{V}_I\end{bmatrix}\qquad \boldsymbol{I}_b=\begin{bmatrix}\boldsymbol{I}_V\\\boldsymbol{I}_C\\\boldsymbol{I}_G\\\boldsymbol{I}_R\\\boldsymbol{I}_L\\\boldsymbol{I}_I\end{bmatrix}$$

现在，设网络的基本割集矩阵为 $\boldsymbol{Q}_f=[\boldsymbol{I},\ \boldsymbol{F}]$. 由假定 (i)—(vi) 可见，将上面确定的 $\boldsymbol{A}$ 变换成标准形后，仅将与电阻相应的列适当交换，便可得到基本割集矩阵. 对应于表 1 中枝的划分，将部分矩阵 $\boldsymbol{F}$ 分解成下列分块矩阵:

$$\boldsymbol{F}=\begin{array}{c}\begin{array}{ccc}\overbrace{\quad}^{n_R} & \overbrace{\quad}^{n_L} & \overbrace{\quad}^{n_I}\end{array}\\ \begin{bmatrix}\boldsymbol{F}_{VR} & \boldsymbol{F}_{VL} & \boldsymbol{F}_{VI}\\ \boldsymbol{F}_{CR} & \boldsymbol{F}_{CL} & \boldsymbol{F}_{CI}\\ \boldsymbol{F}_{GR} & \boldsymbol{F}_{GL} & \boldsymbol{F}_{GI}\end{bmatrix}\end{array}\begin{array}{l}\\ \}\,n_V\\ \}\,n_C\\ \}\,n_G\end{array}$$

这时，基本割集方程式 (14) 和基本回路方程式 (12) 可以分别写成下列形式:

$$\begin{bmatrix}\boldsymbol{I}_V(t)\\\boldsymbol{I}_C(t)\\\boldsymbol{I}_G(t)\end{bmatrix}=\begin{bmatrix}-\boldsymbol{F}_{VR} & -\boldsymbol{F}_{VL} & -\boldsymbol{F}_{VI}\\ -\boldsymbol{F}_{CR} & -\boldsymbol{F}_{CL} & -\boldsymbol{F}_{CI}\\ -\boldsymbol{F}_{GR} & -\boldsymbol{F}_{GL} & -\boldsymbol{F}_{GI}\end{bmatrix}\begin{bmatrix}\boldsymbol{I}_R(t)\\\boldsymbol{I}_L(t)\\\boldsymbol{I}_I(t)\end{bmatrix}\tag{22}$$

$$\begin{bmatrix}\boldsymbol{V}_R(t)\\\boldsymbol{V}_L(t)\\\boldsymbol{V}_I(t)\end{bmatrix}=\begin{bmatrix}\boldsymbol{F}_{VR}^T & \boldsymbol{F}_{CR}^T & \boldsymbol{F}_{GR}^T\\ \boldsymbol{F}_{VL}^T & \boldsymbol{F}_{CL}^T & \boldsymbol{F}_{GL}^T\\ \boldsymbol{F}_{VI}^T & \boldsymbol{F}_{CI}^T & \boldsymbol{F}_{GI}^T\end{bmatrix}\begin{bmatrix}\boldsymbol{V}_V(t)\\\boldsymbol{V}_C(t)\\\boldsymbol{V}_G(t)\end{bmatrix}\tag{23}$$

将 $\boldsymbol{I}_C$ 由 (22) 式，$\boldsymbol{V}_L$ 由 (23) 式提出改写后，得

$$\begin{bmatrix}\boldsymbol{I}_C(t)\\\boldsymbol{V}_L(t)\end{bmatrix}=\begin{bmatrix}\mathbf{0} & -\boldsymbol{F}_{CL}\\ \boldsymbol{F}_{CL}^T & \mathbf{0}\end{bmatrix}\begin{bmatrix}\boldsymbol{V}_C(t)\\\boldsymbol{I}_L(t)\end{bmatrix}+\begin{bmatrix}\mathbf{0} & -\boldsymbol{F}_{CR}\\ \boldsymbol{F}_{GL}^L & \mathbf{0}\end{bmatrix}\begin{bmatrix}\boldsymbol{V}_G(t)\\\boldsymbol{I}_R(t)\end{bmatrix}$$
$$+\begin{bmatrix}\mathbf{0} & -\boldsymbol{F}_{CI}\\ \boldsymbol{F}_{VL}^T & \mathbf{0}\end{bmatrix}\begin{bmatrix}\boldsymbol{V}_V(t)\\\boldsymbol{I}_I(t)\end{bmatrix}\tag{24}$$

因此，若消去 $\boldsymbol{V}_R$ 和 $\boldsymbol{I}_R$，即可得到所要求的混合参数形式的关系式. 因此，在 (22)，(23) 式中，若关于 $\boldsymbol{I}_G$ 和 $\boldsymbol{V}_R$ 的关系式

$$\begin{bmatrix}\boldsymbol{I}_G(t)\\\boldsymbol{V}_R(t)\end{bmatrix}=\begin{bmatrix}\mathbf{0} & -\boldsymbol{F}_{GR}\\ \boldsymbol{F}_{GR}^T & \mathbf{0}\end{bmatrix}\begin{bmatrix}\boldsymbol{V}_G(t)\\\boldsymbol{I}_R(t)\end{bmatrix}+\begin{bmatrix}\mathbf{0} & -\boldsymbol{F}_{GL}\\ \boldsymbol{F}_{CR}^T & \mathbf{0}\end{bmatrix}\begin{bmatrix}\boldsymbol{V}_C(t)\\\boldsymbol{I}_L(t)\end{bmatrix}$$
$$+\begin{bmatrix}\mathbf{0} & -\boldsymbol{F}_{GI}\\ \boldsymbol{F}_{VR}^T & \mathbf{0}\end{bmatrix}\begin{bmatrix}\boldsymbol{V}_V(t)\\\boldsymbol{I}_I(t)\end{bmatrix}\tag{25}$$

成立，则由电阻的端电压电流特性式

$$\begin{bmatrix} \boldsymbol{I}_G(t) \\ \boldsymbol{V}_R(t) \end{bmatrix} = \begin{bmatrix} \boldsymbol{G}_G & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{R}_R \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_G(t) \\ \boldsymbol{I}_R(t) \end{bmatrix} \tag{26}$$

即可求得 $\boldsymbol{V}_G$ 和 $\boldsymbol{I}_R$. 式中 $\boldsymbol{G}_G$ 是以包含在树中电阻的电导值为对角线元素的对角矩阵 $\boldsymbol{R}_R$ 是以包含在树余中电阻的电阻值为对角线元素的对角矩阵. 在正定矩阵中, $\boldsymbol{G}_G$、$\boldsymbol{R}_R$ 都是正则的. 将(25)式代入(26)式,消去 $\boldsymbol{I}_G$ 和 $\boldsymbol{V}_R$ 后,得

$$\begin{bmatrix} \boldsymbol{G}_G & \boldsymbol{F}_{GR} \\ -\boldsymbol{F}_{GR}^T & \boldsymbol{R}_R \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_G(t) \\ \boldsymbol{I}_R(t) \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} & -\boldsymbol{F}_{GL} \\ \boldsymbol{F}_{CR}^T & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_C(t) \\ \boldsymbol{I}_L(t) \end{bmatrix} + \begin{bmatrix} \boldsymbol{0} & -\boldsymbol{F}_{GI} \\ \boldsymbol{F}_{VR}^T & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_V(t) \\ \boldsymbol{I}_I(t) \end{bmatrix} \tag{27}$$

如 B-I-1 中(10)式所示, $[\boldsymbol{V}_G^T, \boldsymbol{I}_R^T]^T$ 的系数矩阵可以表示成 3 个矩阵之积

$$\begin{bmatrix} \boldsymbol{G}_G & \boldsymbol{F}_{GR} \\ -\boldsymbol{F}_{GR}^T & \boldsymbol{R}_R \end{bmatrix} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ -\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1} & \boldsymbol{I} \end{bmatrix} \times \begin{bmatrix} \boldsymbol{G}_G & \boldsymbol{0} \\ \boldsymbol{0} & \mathscr{R} \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & \boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix}$$

$$\mathscr{R} \triangleq \boldsymbol{R}_R + \boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}$$

因为 $\mathscr{R}$ 和 $\boldsymbol{G}_G$ 是正定正则的,故其逆矩阵为

$$\begin{aligned} \begin{bmatrix} \boldsymbol{G}_G & \boldsymbol{F}_{GR} \\ -\boldsymbol{F}_{GR}^T & \boldsymbol{R}_R \end{bmatrix}^{-1} &= \begin{bmatrix} \boldsymbol{I} & -\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR} \\ \boldsymbol{0} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{G}_G^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \mathscr{R}^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1} & \boldsymbol{I} \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{G}_G^{-1} - \boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1} & -\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\mathscr{R}^{-1} \\ \mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1} & \mathscr{R}^{-1} \end{bmatrix}. \end{aligned}$$

因此,由(27)式得

$$\begin{aligned} \begin{bmatrix} \boldsymbol{V}_G(t) \\ \boldsymbol{I}_R(t) \end{bmatrix} &= \begin{bmatrix} -\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{CR}^T & -\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GL} + \boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GL} \\ \mathscr{R}^{-1}\boldsymbol{F}_{CR}^T & -\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GL} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_C(t) \\ \boldsymbol{I}_L(t) \end{bmatrix} \\ &+ \begin{bmatrix} -\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{VR}^T & -\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GI} + \boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GL} \\ \mathscr{R}^{-1}\boldsymbol{F}_{VR}^T & -\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GI} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_V(t) \\ \boldsymbol{I}_I(t) \end{bmatrix} \end{aligned} \tag{28}$$

将(28)式代入(24)式,便得到所要求的混合形关系式

$$\begin{bmatrix} \boldsymbol{I}_C(t) \\ \boldsymbol{V}_L(t) \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_{CC} & \boldsymbol{A}_{CL} \\ \boldsymbol{A}_{LC} & \boldsymbol{A}_{LL} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_C(t) \\ \boldsymbol{I}_L(t) \end{bmatrix} + \begin{bmatrix} \boldsymbol{A}_{CV} & \boldsymbol{A}_{CI} \\ \boldsymbol{A}_{LV} & \boldsymbol{A}_{LI} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_V(t) \\ \boldsymbol{I}_I(t) \end{bmatrix} \tag{29}$$

式中

$$\begin{aligned}
&\boldsymbol{A}_{CC} \triangleq -\boldsymbol{F}_{CR}\mathscr{R}^{-1}\boldsymbol{F}_{CR}^T \\
&\boldsymbol{A}_{CL} \triangleq -\boldsymbol{F}_{CL} + \boldsymbol{F}_{CR}\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GL} \\
&\boldsymbol{A}_{LC} \triangleq \boldsymbol{F}_{CL}^T - \boldsymbol{F}_{GL}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{CR}^T \\
&\boldsymbol{A}_{LL} \triangleq -\boldsymbol{F}_{GL}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GL} + \boldsymbol{F}_{GL}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GL} \\
&\boldsymbol{A}_{CV} \triangleq -\boldsymbol{F}_{CR}\mathscr{R}^{-1}\boldsymbol{F}_{VR}^T \\
&\boldsymbol{A}_{CI} \triangleq -\boldsymbol{F}_{CI} + \boldsymbol{F}_{CR}\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GI} \\
&\boldsymbol{A}_{LV} \triangleq \boldsymbol{F}_{VL}^T - \boldsymbol{F}_{GL}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{VR}^T \\
&\boldsymbol{A}_{LI} \triangleq -\boldsymbol{F}_{GL}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GI} + \boldsymbol{F}_{GL}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GI}
\end{aligned}$$

此外, 在树不包含电阻的情况下, 由上面定义可见, $\boldsymbol{A}_{LL}$ 和 $\boldsymbol{A}_{LI}$ 不确定. 这时, 若回忆一下 (29) 式的推导过程就会明白, $\boldsymbol{A}_{LL}$ 和 $\boldsymbol{A}_{LI}$ 是大小与电感个数及电流源个数相对应的零阵. 同样, 在树余不包含电阻的情况下, $\boldsymbol{A}_{CC}$ 和 $\boldsymbol{A}_{CV}$ 是大小与电容个数及电感个数

相对应的零阵．那么，若令 $\boldsymbol{C}_C$ 为以电容器电容值为对角线上元素的对角矩阵，$\boldsymbol{L}_L$ 为以电感的电感值为主对角线元素的对角矩阵,则电容和电感的端电压-电流特性为

$$\begin{bmatrix} \boldsymbol{I}_C(t) \\ \boldsymbol{V}_L(t) \end{bmatrix} = \begin{bmatrix} \boldsymbol{C}_C & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{L}_L \end{bmatrix} \begin{bmatrix} \dot{\boldsymbol{V}}_C(t) \\ \dot{\boldsymbol{I}}_L(t) \end{bmatrix}. \tag{30}$$

$\boldsymbol{C}_C$, $\boldsymbol{L}_L$ 是正定正则的,因此由(29)、(39)式可以求得

$$\begin{bmatrix} \dot{\boldsymbol{V}}_C(t) \\ \dot{\boldsymbol{I}}_L(t) \end{bmatrix} = \begin{bmatrix} \boldsymbol{C}_C^{-1}\boldsymbol{A}_{CC} & \boldsymbol{C}_C^{-1}\boldsymbol{A}_{CL} \\ \boldsymbol{L}_L^{-1}\boldsymbol{A}_{LC} & \boldsymbol{L}_L^{-1}\boldsymbol{A}_{LL} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_C(t) \\ \boldsymbol{I}_L(t) \end{bmatrix} + \begin{bmatrix} \boldsymbol{C}_C^{-1}\boldsymbol{A}_{CV} & \boldsymbol{C}_C^{-1}\boldsymbol{A}_{CI} \\ \boldsymbol{L}_L^{-1}\boldsymbol{A}_{LV} & \boldsymbol{L}_L^{-1}\boldsymbol{A}_{LI} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_V(t) \\ \boldsymbol{I}_I(t) \end{bmatrix}. \tag{31}$$

上式是关于电容电压 $\boldsymbol{V}_C$ 和电感电流 $\boldsymbol{I}_L$ 的标准形，$\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{u}$ 形式的微分方程式.

但是，是否在所讨论的 RLC 网络中的任意一个元件的端电压、电流都可以由（$\boldsymbol{V}_C$, $\boldsymbol{L}_L$）和输入量（$\boldsymbol{V}_V$, $\boldsymbol{I}_I$）表示?

(28)式已经表明,关于 $\boldsymbol{V}_G$、$\boldsymbol{I}_R$ 是可能的．对于其它变量(电压、电流),将由(28)式求出的 $\boldsymbol{V}_G$, $\boldsymbol{I}_R$ 代入基本割集方程式(22)、基本回路方程式(23)和电阻端电压-电流方程式(26),同样也可以得到用上述变量表示的方程式．其结果如下:

$$\begin{bmatrix} \boldsymbol{V}_V(t) \\ \boldsymbol{V}_C(t) \\ \boldsymbol{V}_G(t) \\ \boldsymbol{V}_R(t) \\ \boldsymbol{V}_L(t) \\ \boldsymbol{V}_I(t) \end{bmatrix}$$

$$= \left[\begin{array}{c:c} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{I} & \boldsymbol{0} \\ -\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{CR}^T & -\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GL}+\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GL} \\ \boldsymbol{R}_R\mathscr{R}^{-1}\boldsymbol{F}_{CR}^T & -\boldsymbol{R}_R\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GL} \\ \boldsymbol{F}_{CL}^T-\boldsymbol{F}_{GL}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{CR}^T & -\boldsymbol{F}_{GL}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GL}+\boldsymbol{F}_{GL}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_U^{-1}\boldsymbol{F}_{GL} \\ \boldsymbol{F}_{CI}^T-\boldsymbol{F}_{GI}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{CR}^T & -\boldsymbol{F}_{GI}^T\boldsymbol{G}_G^T\boldsymbol{F}_{GL}+\boldsymbol{F}_{GI}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GL} \end{array}\right] \begin{bmatrix} \boldsymbol{V}_C(t) \\ \boldsymbol{I}_L(t) \end{bmatrix}$$

$$+ \left[\begin{array}{c:c} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \\ -\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{VR}^T & -\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GI}+\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GI} \\ \boldsymbol{R}_R\mathscr{R}^{-1}\boldsymbol{F}_{VR}^T & -\boldsymbol{R}_R\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GI} \\ \boldsymbol{F}_{VL}^T-\boldsymbol{F}_{GL}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{VR}^T & -\boldsymbol{F}_{GL}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GI}+\boldsymbol{F}_{GL}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GI} \\ \boldsymbol{F}_{VI}^T-\boldsymbol{F}_{GI}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{VR}^T & -\boldsymbol{F}_{GI}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GI}+\boldsymbol{F}_{GI}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{UR}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GI} \end{array}\right] \begin{bmatrix} \boldsymbol{V}_V(t) \\ \boldsymbol{I}_I(t) \end{bmatrix} \tag{32}$$

$$\begin{bmatrix} \boldsymbol{I}_V(t) \\ \boldsymbol{I}_C(t) \\ \boldsymbol{I}_G(t) \\ \boldsymbol{I}_R(t) \\ \boldsymbol{I}_L(t) \\ \boldsymbol{I}_I(t) \end{bmatrix} = \left[\begin{array}{c:c} -\boldsymbol{F}_{VR}\mathscr{R}^{-1}\boldsymbol{F}_{CR}^T & -\boldsymbol{F}_{VL}+\boldsymbol{F}_{VR}\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GL} \\ -\boldsymbol{F}_{CR}\mathscr{R}^{-1}\boldsymbol{F}_{CR}^T & -\boldsymbol{F}_{CL}+\boldsymbol{F}_{CR}\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GL} \\ -\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{CR}^T & -\boldsymbol{F}_{GL}+\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GL} \\ \mathscr{R}^{-1}\boldsymbol{F}_{CR}^T & -\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G{}^{-1}\boldsymbol{F}_{GL} \\ \boldsymbol{0} & \boldsymbol{I} \\ \boldsymbol{0} & \boldsymbol{0} \end{array}\right] \begin{bmatrix} \boldsymbol{V}_C(t) \\ \boldsymbol{I}_L(t) \end{bmatrix}$$

$$
+\left[\begin{array}{c:c}
-F_{VR}\mathscr{R}^{-1}F_{VR}^{T} & -F_{VI}+F_{VR}\mathscr{R}^{-1}F_{GR}^{T}G_{G}^{-1}F_{GI} \\
-F_{CR}\mathscr{R}^{-1}F_{VR}^{T} & -F_{CI}+F_{CR}\mathscr{R}^{-1}F_{GR}^{T}G_{G}^{-1}F_{GI} \\
-F_{GR}\mathscr{R}^{-1}F_{VR}^{T} & -F_{GI}+F_{GR}\mathscr{R}^{-1}F_{GR}^{T}G_{G}^{-1}F_{GI} \\
\mathscr{R}^{-1}F_{VR}^{T} & -\mathscr{R}^{-1}F_{GR}^{T}G_{G}^{-1}F_{GI} \\
\mathbf{0} & \mathbf{0} \\
\mathbf{0} & I
\end{array}\right]\begin{bmatrix} V_V(t) \\ I_I(t) \end{bmatrix} \tag{33}
$$

可见，无论选择网络中的那一个变量(电压、电流)作为输出 $\boldsymbol{y}$，$\boldsymbol{y}$ 都可以用 $\boldsymbol{y}=\boldsymbol{Cx}+\boldsymbol{Du}$ 的形式表示.

结论:

在满足假定 A(i)—(vi) 的 RLC 网络中，电容电压和电感电流通常总可以选作状态变量，将任意元件上的电压、电流作为输出时，标准形状态方程式 $\dot{\boldsymbol{x}}=\boldsymbol{Ax}+\boldsymbol{Bu}$ 和输出方程式 $\boldsymbol{y}=\boldsymbol{Cx}+\boldsymbol{Du}$ 均成立.

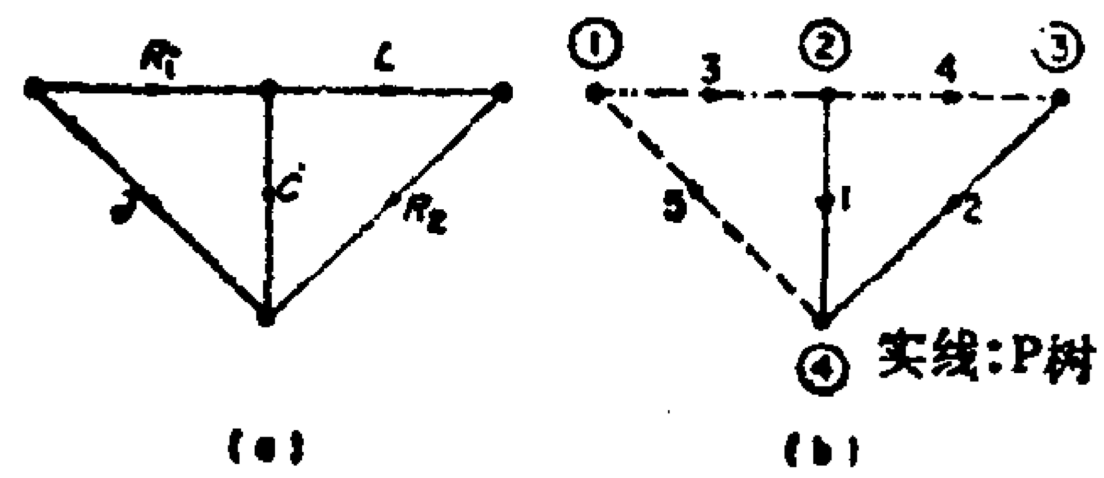


图 13 (a) 图 1 所示电路的图
(b) 它的 P 树

(例 2) 现在用上述方法推导图 1 所示电路的状态方程式. 图 13(a) 是将该电路的枝适当标以方向后的图，在图 13(b) 中，将树选成 P 树，按照电容枝、电阻枝(R_1、R_2 都包含在树中)、电感枝的顺序对枝进行编号. 由该树的基本割集矩阵

$$
\begin{array}{c}
\overbrace{\begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array}}^{\text{枝}} \\
\boldsymbol{Q}_f \triangleq [\boldsymbol{I} \vdots \boldsymbol{F}] = \left[\begin{array}{ccc:cc} 1 & 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{array}\right] \begin{array}{c} c_1 \\ c_2 \\ c_3 \end{array}\Bigg\}\text{割集}
\end{array}
$$

得

$$
\boldsymbol{F} \triangleq \left[\begin{array}{c:c} F_{CL} & F_{CI} \\ \hdashline F_{GL} & F_{GI} \end{array}\right] = \left[\begin{array}{c:c} 1 & -1 \\ \hdashline -1 & 0 \\ 0 & -1 \end{array}\right]
$$

此外，因电容、电感、电阻的端电压-电流特性式分别为

$$
i_1 = C\dot{v}_1(t), \quad v_4(t) = L\dot{i}_4(t),
$$

$$
\begin{bmatrix} i_2(t) \\ i_3(t) \end{bmatrix} = \begin{bmatrix} R_2^{-1} & 0 \\ 0 & R_1^{-1} \end{bmatrix} \begin{bmatrix} v_2(t) \\ v_3(t) \end{bmatrix},
$$

则

$$
\boldsymbol{C}_C = C_1 \quad \boldsymbol{L}_L = L
$$

$$
\boldsymbol{G}_G = \begin{bmatrix} R_2^{-1} & 0 \\ 0 & R_1^{-1} \end{bmatrix}.
$$

将上列各式代入(31)式，便求出标准形状态方程式

$$
\begin{bmatrix} \dot{v}_1(t) \\ \dot{i}_4(t) \end{bmatrix} = \begin{bmatrix} 0 & -\dfrac{1}{C} \\ \dfrac{1}{L} & -\dfrac{R_2}{L} \end{bmatrix} \begin{bmatrix} v_1(t) \\ i_4(t) \end{bmatrix} + \begin{bmatrix} \dfrac{1}{C} \\ 0 \end{bmatrix} i_s(t).
$$

由(32)式可知，这时输出 $y=-v_8$ 可用下式表示

$$y(t)=[1,\ 0]\begin{bmatrix} v_1(t) \\ i_4(t) \end{bmatrix}+R_1 i_8(t).$$

RLCM 网络状态方程式的推导[36]

下面我们来讨论线性非时变 RLCM 网络状态方程式的推导. 根据在 RLC 网络中讲过的同样理由，假定这种 RLCM 网络中不存在仅有电压源的回路和仅有电流源的割集，即

假定: A(i) 不存在仅有电压源的回路;

A(ii) 不存在仅有电流源的割集.

而且，在所讨论的互感(M)中，不包含紧耦合互感.

假定: A(iii) 不存在紧耦合互感.

但是，这里没有提推导 RLC 网络的状态方程式时用过的假定 A(iii)—(vi). 因此，在这里所研究的 RLCM 网络中，有可能存在: 仅具有电容的回路、仅具有电压源和电容的回路、仅具有电感的割集、仅具有电流源和电感的割集. 这时，不能直接把推导 RLC 网络状态方程式时用过的电容电压和电感电流在形式上作为独立变量，必须把电容电压和电感电流中的几个作为因变量. 从图来看，这个事实说明在假定 A(iii)—(vi)不一定成立时不能保证选取 P 树. 因此，引入下面的定义.

相互耦合的电感用图中的枝表示时，忽略其耦合，作为和没有耦合的电感一样处理. 网络用图表示的目的是，将克希霍夫定律用相互独立的方程式表示. 因为克希霍夫定律是只与元件之间的连接关系有关的定律，所以这样处理耦合是正确的. 这里忽略的电感之间的耦合，用(38)式的互感来表示，后者是电感的端电压-电流特性.

[定义 6] 包含所有的电压源枝、尽可能多的电容枝以及尽可能少的电感枝，而不包含电流源枝的树，叫做规范树1)(normal tree, 以下简写成 N 树). 可以认为，N 树中电容枝电压和对应于该树的树余中电感枝的电流是相互独立的. 树余中电容枝电压从属于树中电压源和电容枝的电压，树中电感枝的电流从属于树余中电流源及电感枝的电流. 即假定 A(i), (ii), (iv)成立. 但是在 A(iii)—(iv)不一定成立的 RLCM 网络里，树中电容枝电压与其树余中电感枝电流是在形式上作为独立变量处理的.

现在，我们来讲一下 RLCM 网络状态方程式的推导方法. 首先找出所给网络图的 N

表 2 枝的分类 II

枝的种类	电压向量	电流向量	枝的数量	
电压源枝	$\boldsymbol{V}_V$	$\boldsymbol{I}_V$	n_V	树
树中电容枝	$\boldsymbol{V}_C$	$\boldsymbol{I}_C$	n_C	
树中电阻枝	$\boldsymbol{V}_G$	$\boldsymbol{I}_G$	n_G	
树中电感枝	$\boldsymbol{V}_\Gamma$	$\boldsymbol{I}_\Gamma$	n_Γ	
树余中电容枝	$\boldsymbol{V}_S$	$\boldsymbol{I}_S$	n_S	树余
树余中电阻枝	$\boldsymbol{V}_R$	$\boldsymbol{I}_R$	n_R	
树余中电感枝	$\boldsymbol{V}_L$	$\boldsymbol{I}_L$	n_L	
电流源枝	$\boldsymbol{V}_I$	$\boldsymbol{I}_I$	n_I	

1) 在满足假定 A(i)—(vii)的网络中，N 树和 P 树是等价的. 和 P 树一样，N 树也不是唯一确定的.

树，然后以该 N 树为基础，按着电压源枝、树中电容枝、树中电阻枝、树中电感枝、树余中电容枝、树余中电阻枝、树余中电感枝、电流源枝的顺序对枝进行编号．如果以各枝电压、电流为分量的列向量和枝数采用表 2 中的符号表示，则全枝电压列向量 $\boldsymbol{V}_b$、电流列向量 $\boldsymbol{I}_b$ 可以表示为

$$\boldsymbol{V}_b=\begin{bmatrix} \boldsymbol{V}_V \\ \boldsymbol{V}_C \\ \boldsymbol{V}_G \\ \boldsymbol{V}_\Gamma \\ \boldsymbol{V}_S \\ \boldsymbol{V}_R \\ \boldsymbol{V}_L \\ \boldsymbol{V}_I \end{bmatrix}, \qquad \boldsymbol{I}_b=\begin{bmatrix} \boldsymbol{I}_V \\ \boldsymbol{I}_C \\ \boldsymbol{I}_G \\ \boldsymbol{I}_\Gamma \\ \boldsymbol{I}_S \\ \boldsymbol{I}_R \\ \boldsymbol{I}_L \\ \boldsymbol{I}_I \end{bmatrix}.$$

与此相对应，基本割集矩阵 $\boldsymbol{Q}_f=[\boldsymbol{I}, \boldsymbol{F}]$ 的部分矩阵 $\boldsymbol{F}$ 可以分解成如下分块矩阵

$$\boldsymbol{F}=\begin{array}{c} \begin{array}{cccc} \overbrace{\quad}^{n_S} & \overbrace{\quad}^{n_R} & \overbrace{\quad}^{n_L} & \overbrace{\quad}^{n_I} \end{array} \\ \begin{bmatrix} \boldsymbol{F}_{VS} & \boldsymbol{F}_{VR} & \boldsymbol{F}_{VL} & \boldsymbol{F}_{VI} \\ \boldsymbol{F}_{CS} & \boldsymbol{F}_{CR} & \boldsymbol{F}_{CL} & \boldsymbol{F}_{CI} \\ \boldsymbol{F}_{GS} & \boldsymbol{F}_{GR} & \boldsymbol{F}_{GL} & \boldsymbol{F}_{GI} \\ \boldsymbol{F}_{\Gamma S} & \boldsymbol{F}_{\Gamma R} & \boldsymbol{F}_{\Gamma L} & \boldsymbol{F}_{\Gamma I} \end{bmatrix} \end{array} \begin{array}{l} \\ \}\, n_V \\ \}\, n_C \\ \}\, n_G \\ \}\, n_\Gamma \end{array}$$

因树选为 N 树，故树中电感枝仅与树余中电流源枝及电感枝构成割集，而树余中电容枝仅与树中电压源枝及电容枝构成回路，则

$$\boldsymbol{F}_{\Gamma S}=0, \quad \boldsymbol{F}_{\Gamma R}=0, \quad \boldsymbol{F}_{GS}=0$$

而且，基本割集方程可以分别表示成下列形式

$$\begin{bmatrix} \boldsymbol{I}_V(t) \\ \boldsymbol{I}_C(t) \\ \boldsymbol{I}_G(t) \\ \boldsymbol{I}_\Gamma(t) \end{bmatrix}=\begin{bmatrix} -\boldsymbol{F}_{VS} & -\boldsymbol{F}_{VR} & -\boldsymbol{F}_{VL} & -\boldsymbol{F}_{VI} \\ -\boldsymbol{F}_{CS} & -\boldsymbol{F}_{CR} & -\boldsymbol{F}_{CL} & -\boldsymbol{F}_{CI} \\ \boldsymbol{0} & -\boldsymbol{F}_{GR} & -\boldsymbol{F}_{GL} & -\boldsymbol{F}_{GI} \\ \boldsymbol{0} & \boldsymbol{0} & -\boldsymbol{F}_{\Gamma L} & -\boldsymbol{F}_{\Gamma I} \end{bmatrix}\begin{bmatrix} \boldsymbol{I}_S(t) \\ \boldsymbol{I}_R(t) \\ \boldsymbol{I}_L(t) \\ \boldsymbol{I}_I(t) \end{bmatrix} \tag{34}$$

$$\begin{bmatrix} \boldsymbol{V}_S(t) \\ \boldsymbol{V}_R(t) \\ \boldsymbol{V}_L(t) \\ \boldsymbol{V}_I(t) \end{bmatrix}=\begin{bmatrix} \boldsymbol{F}_{VS}^T & \boldsymbol{F}_{CS}^T & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{F}_{VR}^T & \boldsymbol{F}_{CR}^T & \boldsymbol{F}_{GR}^T & \boldsymbol{0} \\ \boldsymbol{F}_{VL}^T & \boldsymbol{F}_{CL}^T & \boldsymbol{F}_{GL}^T & \boldsymbol{F}_{\Gamma L}^T \\ \boldsymbol{F}_{VI}^T & \boldsymbol{F}_{CI}^T & \boldsymbol{F}_{GI}^T & \boldsymbol{F}_{\Gamma I}^T \end{bmatrix}\begin{bmatrix} \boldsymbol{V}_V(t) \\ \boldsymbol{V}_C(t) \\ \boldsymbol{V}_G(t) \\ \boldsymbol{V}_\Gamma(t) \end{bmatrix} \tag{35}$$

此外，电阻、电容、电感的端电压-电流特性可以分别表示成下列形式

$$\begin{bmatrix} \boldsymbol{I}_G(t) \\ \boldsymbol{V}_R(t) \end{bmatrix}=\begin{bmatrix} \boldsymbol{G}_G & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{R}_R \end{bmatrix}\begin{bmatrix} \boldsymbol{V}_G(t) \\ \boldsymbol{I}_R(t) \end{bmatrix} \tag{36}$$

$$\begin{bmatrix} \boldsymbol{I}_C(t) \\ \boldsymbol{I}_S(t) \end{bmatrix}=\frac{d}{dt}\left\{\begin{bmatrix} \boldsymbol{C}_C & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C}_S \end{bmatrix}\begin{bmatrix} \boldsymbol{V}_C(t) \\ \boldsymbol{V}_S(t) \end{bmatrix}\right\} \tag{37}$$

$$\begin{aligned} \begin{bmatrix} \boldsymbol{V}_\Gamma(t) \\ \boldsymbol{V}_L(t) \end{bmatrix} &=\frac{d}{dt}\left\{\begin{bmatrix} \boldsymbol{L}_{\Gamma\Gamma} & \boldsymbol{L}_{\Gamma L} \\ \boldsymbol{L}_{L\Gamma} & \boldsymbol{L}_{LL} \end{bmatrix}\begin{bmatrix} \boldsymbol{I}_\Gamma(t) \\ \boldsymbol{I}_L(t) \end{bmatrix}\right\} \\ &\triangleq \frac{d}{dt}\left\{\boldsymbol{L}\begin{bmatrix} \boldsymbol{I}_\Gamma(t) \\ \boldsymbol{I}_L(t) \end{bmatrix}\right\} \end{aligned} \tag{38}$$

式中 $\boldsymbol{G}_G$、$\boldsymbol{R}_R$、$\boldsymbol{C}_C$、$\boldsymbol{C}_S$ 是正定的对角矩阵，$\boldsymbol{L}$ 是对称矩阵，其对角线上元素表示自感，非对角线上元素表示互感．根据假定 A(vii)，由于在所讨论的电路中没有紧耦合电感，所以 $\boldsymbol{L}$ 是正定的．

利用(37)、(38)式，消去(34)、(35)式中的 $\boldsymbol{I}_S$、$\boldsymbol{I}_\Gamma$、$\boldsymbol{V}_S$、$\boldsymbol{V}_\Gamma$，便得到和 RLC 网络情况下的(24)式相类似的关系式

$$\frac{d}{dt}\begin{bmatrix} \mathscr{C}\boldsymbol{V}_C(t)+\boldsymbol{F}_{CS}\boldsymbol{C}_S\boldsymbol{F}_{VS}^T\boldsymbol{V}_V(t) \\ \mathscr{L}\boldsymbol{I}_L(t)+(\boldsymbol{F}_{\Gamma L}^T\boldsymbol{L}_{\Gamma\Gamma}-\boldsymbol{L}_{L\Gamma})\boldsymbol{F}_{\Gamma I}\boldsymbol{I}_I(t) \end{bmatrix}=\begin{bmatrix} \boldsymbol{0} & -\boldsymbol{F}_{CR} \\ \boldsymbol{F}_{GL}^T & \boldsymbol{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{V}_G(t) \\ \boldsymbol{I}_R(t) \end{bmatrix}$$

$$+\begin{bmatrix} \boldsymbol{0} & -\boldsymbol{F}_{CL} \\ \boldsymbol{F}_{CL}^T & \boldsymbol{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{V}_C(t) \\ \boldsymbol{I}_L(t) \end{bmatrix}+\begin{bmatrix} \boldsymbol{0} & -\boldsymbol{F}_{CI} \\ \boldsymbol{F}_{VL}^T & \boldsymbol{0} \end{bmatrix}\begin{bmatrix} \boldsymbol{V}_V(t) \\ \boldsymbol{I}_I(t) \end{bmatrix} \tag{39}$$

$$\mathscr{C}\triangleq\boldsymbol{C}_C+\boldsymbol{F}_{CS}\boldsymbol{C}_S\boldsymbol{F}_{CS}^T$$

$$\mathscr{L}\triangleq[-\boldsymbol{F}_{\Gamma L}^T\boldsymbol{I}]\begin{bmatrix} \boldsymbol{L}_{\Gamma\Gamma} & \boldsymbol{L}_{\Gamma L} \\ \boldsymbol{L}_{L\Gamma} & \boldsymbol{L}_{LL} \end{bmatrix}\begin{bmatrix} -\boldsymbol{F}_{\Gamma L} \\ \boldsymbol{I} \end{bmatrix}=[-\boldsymbol{F}_{\Gamma L}^T\boldsymbol{I}]\boldsymbol{L}\begin{bmatrix} -\boldsymbol{F}_{\Gamma L} \\ \boldsymbol{I} \end{bmatrix}.$$

根据正定矩阵的定义，因 $\mathscr{C}$ 是正定矩阵 $\boldsymbol{C}_C$ 和准正定矩阵 $\boldsymbol{F}_{CS}\boldsymbol{C}_S\boldsymbol{F}_{CS}^T$[1)]之和，故 $\mathscr{C}$ 是正定的．因 $\boldsymbol{L}$ 是正定的，所以除了零列向量以外，对于任意 n_L 维列向量 $\boldsymbol{x}$，有

$$\begin{bmatrix} -\boldsymbol{F}_{\Gamma L} \\ \boldsymbol{I} \end{bmatrix}\boldsymbol{x}=\begin{bmatrix} -\boldsymbol{F}_{\Gamma L}\boldsymbol{x} \\ \boldsymbol{x} \end{bmatrix}\neq\boldsymbol{0},$$

则 $\mathscr{L}$ 是正定的[2)]．因此，$\mathscr{C}$、$\mathscr{L}$ 是正则的[3)]．

将(34)式中 $\boldsymbol{I}_G$ 的关系式和(35)式中 $\boldsymbol{V}_R$ 的关系式代入(36)式，然后对 $\boldsymbol{V}_G$、$\boldsymbol{I}_R$ 求解，便得到和 RLC 网络情况下完全相同的(28)式，将其代入(39)式，得

$$\frac{d}{dt}\begin{bmatrix} \mathscr{C}\boldsymbol{V}_C(t)+\boldsymbol{F}_{CS}\boldsymbol{C}_S\boldsymbol{F}_{VS}^T\boldsymbol{V}_V(t) \\ \mathscr{L}\boldsymbol{I}_L(t)+(\boldsymbol{F}_{\Gamma L}^T\boldsymbol{L}_{\Gamma\Gamma}-\boldsymbol{L}_{L\Gamma})\boldsymbol{F}_{\Gamma I}\boldsymbol{I}_I(t) \end{bmatrix}$$

$$=\begin{bmatrix} \boldsymbol{A}_{CC} & \boldsymbol{A}_{CL} \\ \boldsymbol{A}_{LC} & \boldsymbol{A}_{LL} \end{bmatrix}\begin{bmatrix} \boldsymbol{V}_C(t) \\ \boldsymbol{I}_L(t) \end{bmatrix}+\begin{bmatrix} \boldsymbol{A}_{CV} & \boldsymbol{A}_{CI} \\ \boldsymbol{A}_{LV} & \boldsymbol{A}_{LI} \end{bmatrix}\begin{bmatrix} \boldsymbol{V}_V(t) \\ \boldsymbol{I}_I(t) \end{bmatrix}.$$

在这里若引入新的变量

$$\left.\begin{aligned} \boldsymbol{q}(t)&\triangleq\mathscr{C}\boldsymbol{V}_C(t)+\boldsymbol{F}_{CS}\boldsymbol{C}_S\boldsymbol{F}_{VS}^T\boldsymbol{V}_V(t) \\ \boldsymbol{\phi}(t)&\triangleq\mathscr{L}\boldsymbol{I}_L(t)+(\boldsymbol{F}_{\Gamma L}^T\boldsymbol{L}_{\Gamma\Gamma}-\boldsymbol{L}_{L\Gamma})\boldsymbol{F}_{\Gamma I}\boldsymbol{I}_I(t) \end{aligned}\right\} \tag{40}$$

便得到关于 $\boldsymbol{q}$，$\boldsymbol{\phi}$ 的标准形微分方程式

$$\begin{bmatrix} \dot{\boldsymbol{q}}(t) \\ \dot{\boldsymbol{\phi}}(t) \end{bmatrix}=\begin{bmatrix} \boldsymbol{A}_{CC}\mathscr{C}^{-1} & \boldsymbol{A}_{CL}\mathscr{L}^{-1} \\ \boldsymbol{A}_{LC}\mathscr{C}^{-1} & \boldsymbol{A}_{LL}\mathscr{L}^{-1} \end{bmatrix}\begin{bmatrix} \boldsymbol{q}(t) \\ \boldsymbol{\phi}(t) \end{bmatrix}$$

$$+\begin{bmatrix} \boldsymbol{A}_{CV}-\boldsymbol{A}_{CC}\mathscr{C}^{-1}\boldsymbol{F}_{CS}\boldsymbol{C}_S\boldsymbol{F}_{VS}^T & \boldsymbol{A}_{CI}-\boldsymbol{A}_{CL}\mathscr{L}^{-1}(\boldsymbol{F}_{\Gamma L}^T\boldsymbol{L}_{\Gamma\Gamma}-\boldsymbol{L}_{L\Gamma})\boldsymbol{F}_{\Gamma I} \\ \boldsymbol{A}_{LV}-\boldsymbol{A}_{LC}\mathscr{C}^{-1}\boldsymbol{F}_{CS}\boldsymbol{C}_S\boldsymbol{F}_{VS}^T & \boldsymbol{A}_{LI}-\boldsymbol{A}_{LL}\mathscr{L}^{-1}(\boldsymbol{F}_{\Gamma L}^T\boldsymbol{L}_{\Gamma\Gamma}-\boldsymbol{L}_{L\Gamma})\boldsymbol{F}_{\Gamma I} \end{bmatrix}\begin{bmatrix} \boldsymbol{V}_V(t) \\ \boldsymbol{I}_I(t) \end{bmatrix}. \tag{41}$$

---

1) 若 $n\times n$ 阶对称矩阵 $\boldsymbol{A}\in R^{n\times n}$ 是正定的（参看 P-I-1），则由 $\boldsymbol{A}$ 和任意 $\boldsymbol{B}\in R^{n\times m}$ 组成的 $\boldsymbol{B}^T\boldsymbol{A}\boldsymbol{B}\in R^{m\times m}$ 是准正定的．特别是，对于除了零向量以外的任意 $\boldsymbol{x}\in R^{m\times 1}$，若 $\boldsymbol{B}\boldsymbol{x}\neq\boldsymbol{0}$，则 $\boldsymbol{B}^T\boldsymbol{A}\boldsymbol{B}$ 是正定的．其证明如下：设 $\boldsymbol{y}\triangleq\boldsymbol{B}\boldsymbol{x}\in R^{n\times 1}$，则 $\boldsymbol{x}^T\boldsymbol{B}^T\boldsymbol{A}\boldsymbol{B}\boldsymbol{x}=\boldsymbol{y}^T\boldsymbol{A}\boldsymbol{y}$ 成立．因为 $\boldsymbol{A}$ 是正定的，若 $\boldsymbol{y}\neq\boldsymbol{0}$，则 $\boldsymbol{y}^T\boldsymbol{A}\boldsymbol{y}>0$；若 $\boldsymbol{y}=0$，则 $\boldsymbol{y}^T\boldsymbol{A}\boldsymbol{y}=\boldsymbol{0}$，则无论如何 $\boldsymbol{y}^T\boldsymbol{A}\boldsymbol{y}=\boldsymbol{x}^T\boldsymbol{B}^T\boldsymbol{A}\boldsymbol{B}\boldsymbol{x}$ 不会是负的．因此，$\boldsymbol{B}^T\boldsymbol{A}\boldsymbol{B}$ 是准正定的．在这里不能断定 $\boldsymbol{B}^T\boldsymbol{A}\boldsymbol{B}$ 是正定的，因为一般即使 $\boldsymbol{x}\neq\boldsymbol{0}$，也可能 $\boldsymbol{y}=\boldsymbol{B}\boldsymbol{x}=\boldsymbol{0}$．如果仅限于 $\boldsymbol{x}=\boldsymbol{0}$ 时才有 $\boldsymbol{B}\boldsymbol{x}=\boldsymbol{0}$，则对于任意 $\boldsymbol{x}\neq\boldsymbol{0}$ 有 $\boldsymbol{y}^T\boldsymbol{A}\boldsymbol{y}=\boldsymbol{x}^T\boldsymbol{B}^T\boldsymbol{A}\boldsymbol{B}\boldsymbol{x}>0$，即 $\boldsymbol{B}^T\boldsymbol{A}\boldsymbol{B}$ 是正定的．

2) 同上．

3) 正定矩阵 $\boldsymbol{A}$ 是正则的．其证明如下：设 $\boldsymbol{A}$ 是正定但不是正则的，则线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 具有除零列向量以外的解（参看 B-I-4）．设该解为 $\boldsymbol{x}_a(\neq 0)$，则对于 $\boldsymbol{x}_a$ 有 $\boldsymbol{x}_a^T\boldsymbol{A}\boldsymbol{x}_a=\boldsymbol{0}$，这和 $\boldsymbol{A}$ 是正定的假定相矛盾．因此，正定矩阵 $\boldsymbol{A}$ 是正则的．

若代替 $\boldsymbol{q}$、$\boldsymbol{\phi}$，引入分别具有电压、电流因次的变量

$$\left.\begin{aligned}\boldsymbol{e}(t) &\triangleq \boldsymbol{V}_C(t)+\mathscr{C}^{-1}\boldsymbol{F}_{CS}\boldsymbol{C}_S\boldsymbol{F}_{VS}^T\boldsymbol{V}_V(t)=\mathscr{C}^{-1}\boldsymbol{q}(t)\\ \boldsymbol{j}(t) &\triangleq \boldsymbol{I}_L(t)+\mathscr{L}^{-1}(\boldsymbol{F}_{\Gamma L}^T\boldsymbol{L}_{\Gamma\Gamma}-\boldsymbol{L}_{\Gamma L})\boldsymbol{F}_{\Gamma I}\boldsymbol{I}_I(t)=\mathscr{L}^{-1}\boldsymbol{\phi}(t)\end{aligned}\right\} \tag{40$'$}$$

也可以得到关于 $\boldsymbol{e}$、$\boldsymbol{j}$ 的标准形微分方程式．因此，上述关于 $\boldsymbol{q}$、$\boldsymbol{\phi}$ 的讨论，对于 $\boldsymbol{e}$、$\boldsymbol{j}$ 也几乎完全成立．但是，即使对于时变线性 R、L、C、M 网络 $[R(t)L(t)C(t)M(t)$ 网络$]$，若引入 $\boldsymbol{q}$、$\boldsymbol{\phi}$，(41)式的标准形微分方程式也照样成立．因上述 RLCM 网络状态方程式的推导方法很容易推广到 $R(t)L(t)C(t)M(t)$ 网络，所以下面将使用 $\boldsymbol{q}$、$\boldsymbol{\phi}$．当仅限于讨论RLCM 网络时，也可以使用 $\boldsymbol{e}$、$\boldsymbol{j}$，它很容易和满足假定 A(i)—A(vi)的 RLC 网络情况下的状态变量 $\boldsymbol{V}_C$，$\boldsymbol{I}_L$ 相比较．

那末，在所讨论的网络中，令变量 $\boldsymbol{x}=[\boldsymbol{q}^T,\ \boldsymbol{\phi}^T]^T$ 和输入 $\boldsymbol{u}=[\boldsymbol{V}_V^T,\ \boldsymbol{I}_I^T]^T$．若一切变量(电压、电流)都可以用 $\boldsymbol{y}=\boldsymbol{Cx}+\boldsymbol{Du}$ 的形式表示，则$[\boldsymbol{q}^T,\ \boldsymbol{\phi}^T]^T$ 可称作状态变量，而与将那个变量选为输出 $\boldsymbol{y}$ 无关．但是实际上，当 $\boldsymbol{y}$ 任意选定时，不一定表示成 $\boldsymbol{y}=\boldsymbol{Cx}+\boldsymbol{Du}$ 的形式，有时必须用输入量微分表示．例如，电压源电流 $\boldsymbol{I}_V$ 和电流源电压 $\boldsymbol{V}_I$ 可以表示如下：

$$\begin{aligned}\boldsymbol{I}_V(t)=&\{-\boldsymbol{F}_{VR}\mathscr{R}^{-1}\boldsymbol{F}_{CR}^T-\boldsymbol{F}_{VS}\boldsymbol{C}_S\boldsymbol{F}_{CS}^T\mathscr{C}^{-1}\boldsymbol{A}_{CC}\}\mathscr{C}^{-1}\boldsymbol{q}(t)+\{-\boldsymbol{F}_{VL}+\boldsymbol{F}_{VR}\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GL}\\ &-\boldsymbol{F}_{VS}\boldsymbol{C}_S\boldsymbol{F}_{CS}^T\mathscr{C}^{-1}\boldsymbol{A}_{CL}\}\mathscr{L}^{-1}\boldsymbol{\phi}(t)+\{-\boldsymbol{F}_{VR}\mathscr{R}^{-1}\boldsymbol{F}_{VR}^T-\boldsymbol{F}_{VS}\boldsymbol{C}_S\boldsymbol{F}_{CS}^T\boldsymbol{A}_{CV}\\ &+(\boldsymbol{F}_{VR}\mathscr{R}^{-1}\boldsymbol{F}_{CR}^T+\boldsymbol{F}_{VS}\boldsymbol{C}_S\boldsymbol{F}_{CS}^T\boldsymbol{A}_{CC})\mathscr{C}^{-1}\boldsymbol{F}_{CS}\boldsymbol{C}_S\boldsymbol{F}_{VS}^T\}\boldsymbol{V}_V(t)+\{-\boldsymbol{F}_{VI}\\ &+\boldsymbol{F}_{VR}\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GI}+\boldsymbol{F}_{VS}\boldsymbol{C}_S\boldsymbol{F}_{CS}^T\mathscr{C}^{-1}\boldsymbol{A}_{CI}+(\boldsymbol{F}_{VL}-\boldsymbol{F}_{VR}\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GL}\\ &-\boldsymbol{F}_{VS}\boldsymbol{C}_S\boldsymbol{F}_{CS}^T\mathscr{C}^{-1}\boldsymbol{A}_{CL})\times\mathscr{L}^{-1}(\boldsymbol{F}_{\Gamma L}^T\boldsymbol{L}_{\Gamma\Gamma}-\boldsymbol{L}_{L\Gamma})\boldsymbol{F}_{\Gamma I}\}\boldsymbol{I}_I(t)+\{-\boldsymbol{F}_{VS}\boldsymbol{C}_S\boldsymbol{F}_{VS}^T\\ &+\boldsymbol{F}_{VS}\boldsymbol{C}_S\boldsymbol{F}_{CS}^T\mathscr{C}^{-1}\boldsymbol{F}_{CS}\boldsymbol{C}_S\boldsymbol{F}_{VS}^T\}\dot{\boldsymbol{V}}_V(t).\end{aligned}\tag{42}$$

$$\begin{aligned}\boldsymbol{V}_I(t)=&\{\boldsymbol{F}_{CI}^T-\boldsymbol{F}_{\Gamma I}^T(\boldsymbol{L}_{\Gamma\Gamma}\boldsymbol{F}_{\Gamma L}-\boldsymbol{L}_{\Gamma L})\mathscr{L}^{-1}\boldsymbol{A}_{LC}-\boldsymbol{F}_{GI}^T(\boldsymbol{G}_G^{-1}-\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1})\\ &\times\boldsymbol{F}_{GR}\boldsymbol{R}_R^{-1}\boldsymbol{F}_{CR}^T\}\mathscr{C}^{-1}\boldsymbol{q}(t)+\{-\boldsymbol{F}_{\Gamma I}^T(\boldsymbol{L}_{\Gamma\Gamma}\boldsymbol{F}_{\Gamma L}-\boldsymbol{L}_{\Gamma L})\mathscr{L}^{-1}\boldsymbol{A}_{LL}-\boldsymbol{F}_{GI}(\boldsymbol{G}_G^{-1}\\ &-\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1})\boldsymbol{F}_{GL}\}\mathscr{L}^{-1}\boldsymbol{\phi}(t)+\{\boldsymbol{F}_{VI}^T-\boldsymbol{F}_{CI}^T\mathscr{C}^{-1}\boldsymbol{F}_{CS}\boldsymbol{C}_S\boldsymbol{F}_{VS}^T\\ &-\boldsymbol{F}_{\Gamma I}^T(\boldsymbol{L}_{\Gamma\Gamma}\boldsymbol{F}_{\Gamma L}-\boldsymbol{L}_{\Gamma L})\mathscr{L}^{-1}(\boldsymbol{A}_{LV}-\boldsymbol{A}_{LC}\mathscr{C}^{-1}\boldsymbol{F}_{CS}\boldsymbol{C}_S\boldsymbol{F}_{VS}^T)-\boldsymbol{F}_{GI}^T(\boldsymbol{G}_G^{-1}\\ &-\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1})\boldsymbol{F}_{GR}\boldsymbol{R}_R^{-1}(\boldsymbol{F}_{VR}^T-\boldsymbol{F}_{CR}^T\mathscr{C}^{-1}\boldsymbol{F}_{CS}\boldsymbol{C}_S\boldsymbol{F}_{VS}^T)\}\boldsymbol{V}_V(t)\\ &+\{-\boldsymbol{F}_{GI}^T(\boldsymbol{G}_G^{-1}-\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\mathscr{R}^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{G}_G^{-1})(\boldsymbol{F}_{GI}-\boldsymbol{F}_{GL}\mathscr{L}^{-1}(\boldsymbol{F}_{\Gamma L}^T\boldsymbol{L}_{\Gamma\Gamma}-\boldsymbol{L}_{L\Gamma})\boldsymbol{F}_{\Gamma I})\\ &+\boldsymbol{F}_{\Gamma I}^T(\boldsymbol{L}_{\Gamma\Gamma}\boldsymbol{F}_{\Gamma L}-\boldsymbol{L}_{\Gamma L})\mathscr{L}^{-1}(\boldsymbol{A}_{LI}-\boldsymbol{A}_{LL}\mathscr{L}^{-1}(\boldsymbol{F}_{L\Gamma}^T\boldsymbol{L}_{\Gamma\Gamma}-\boldsymbol{L}_{L\Gamma})\boldsymbol{F}_{\Gamma I})\}\boldsymbol{I}_I(t)\\ &+\{-\boldsymbol{F}_{\Gamma I}^T\boldsymbol{L}_{\Gamma\Gamma}\boldsymbol{F}_{\Gamma I}+\boldsymbol{F}_{\Gamma I}^T(\boldsymbol{L}_{\Gamma\Gamma}\boldsymbol{F}_{\Gamma L}-\boldsymbol{L}_{\Gamma L})\mathscr{L}^{-1}(\boldsymbol{F}_{\Gamma L}^T\boldsymbol{L}_{\Gamma\Gamma}-\boldsymbol{L}_{L\Gamma})\boldsymbol{F}_{\Gamma I}\}\dot{\boldsymbol{I}}_I(t).\end{aligned}\tag{43}$$

即若利用 $\boldsymbol{x}=[\boldsymbol{q}^T,\ \boldsymbol{\phi}^T]^T$，一般可以推导出下列形式的输入输出关系式：

$$\begin{aligned}\dot{\boldsymbol{x}}&=\boldsymbol{Ax}+\boldsymbol{Bu}\\ \boldsymbol{y}&=\boldsymbol{Cx}+\boldsymbol{Du}+\boldsymbol{E}\dot{\boldsymbol{u}}.\end{aligned}$$

可以说，这里将输入量微分 $\dot{\boldsymbol{u}}$ 看成新的输入量，$[\boldsymbol{q}^T,\ \boldsymbol{\phi}^T]^T$ 是状态变量(与那一个变量选为 $\boldsymbol{y}$ 无关)．为了区别于满足 A-I-1 中条件(P.2)、(P.3)的状态变量，称$[\boldsymbol{q}^T,\ \boldsymbol{\phi}^T]^T$ 为准状态变量(pseudo state variable)．

在所讨论的网络中，若假定 A(iii)—A(vi)成立，则矩阵 $\boldsymbol{F}$ 的分块矩阵中 $\boldsymbol{F}_{VS}$、$\boldsymbol{F}_{CS}$、$\boldsymbol{F}_{\Gamma L}$、$\boldsymbol{F}_{\Gamma L}$ 不存在．这时 $\boldsymbol{q}$、$\boldsymbol{\phi}$ 分别是电容器上的电荷及电感线圈中的磁通一样的物理变量．而且由(42)式的 $\boldsymbol{I}_V$，(43)式的 $\boldsymbol{V}_I$ 可见，无论将网络的那一个变量(电压、电流)作为输出量 $\boldsymbol{y}$，都可以表示成 $\boldsymbol{y}=\boldsymbol{Cx}+\boldsymbol{Du}$ 的形式，即$[\boldsymbol{q}^T,\ \boldsymbol{\phi}^T]^T$ 满足状态变量的条件(P.2)、(P.3)．因为这里所讨论的是线性非时变网络，和满足假定 A(i)—A(vi)的 RLC 网络的

情况一样，当然也可以不用 $\boldsymbol{q}$、$\boldsymbol{\phi}$，而用电容器上电压 $\boldsymbol{V}_C=\boldsymbol{C}_C^{-1}\boldsymbol{q}$，电感线圈中的电流 $\boldsymbol{I}_L=\boldsymbol{L}_{LL}^{-1}\boldsymbol{\phi}$ 作为状态变量.

归纳以上，得出下列结论：

在假定 A(i)、(ii)、(vii) 成立的非时变线性网络中，对于用 (40) 式定义的变量 $\boldsymbol{x}=[\boldsymbol{q}^T, \boldsymbol{\phi}^T]^T$，标准形微分方程式成立．一般，输出(电压、电流)可以用 $\boldsymbol{q}$、$\boldsymbol{\phi}$ 和输入量及输入量微分的线性组合表示．若假定 A(iii)—$\boldsymbol{A}$(vi) 成立，则输出量仅由 $\boldsymbol{q}$、$\boldsymbol{\phi}$ 和输入量表示，这时 $\boldsymbol{q}$ 是电容器上的电荷，$\boldsymbol{\phi}$ 是电感线圈中的磁通.

$\boldsymbol{R}(t)$, $\boldsymbol{L}(t)$, $\boldsymbol{C}(t)$, $\boldsymbol{M}(t)$ 网络的状态方程式[36,37]

现在讨论网络中的电阻、电感、电容、互感的特性均随时间变化的时变 $\boldsymbol{R}(t)$、$\boldsymbol{L}(t)$、$C(t)$、$M(t)$ 网络．因为推导非时变线性网络状态方程式的思路和方法原则上也完全适用于这种网络，所以这里只作简单说明.

首先，设假定条件 A(i)，(ii) 成立．根据 (40) 式 $\boldsymbol{q}$、$\boldsymbol{\phi}$ 的定义，若式中 $\boldsymbol{G}_G(t)$、$\mathscr{R}(t)$、$\mathscr{C}(t)$、$\mathscr{L}(t)$ 对所有的 t 均为正则的，则 $\boldsymbol{G}_G^{-1}(t)$、$\mathscr{R}^{-1}(t)$、$\mathscr{C}^{-1}(t)$、$\mathscr{L}^{-1}(t)$ 存在，关于 $[\boldsymbol{q}^T, \boldsymbol{\phi}^T]^T$ 的标准形微分方程式 (41)，即

$$\dot{\boldsymbol{x}}(t)=\boldsymbol{A}(t)\boldsymbol{x}(t)+\boldsymbol{B}(t)\boldsymbol{u}(t)$$

形的微分方程式成立．一般．输出都用 $\boldsymbol{q}$、$\boldsymbol{\phi}$，输入量及输入量微分的线性组合表示．例如，和以前一样，树余中电容器枝的电压、电流可以分别用下式表示

$$\boldsymbol{V}_S(t)=\boldsymbol{F}_{CS}^T\boldsymbol{C}^{-1}(t)\boldsymbol{q}(t)+(\boldsymbol{F}_{VS}^T-\boldsymbol{F}_{CS}^T\mathscr{C}^{-1}(t)\boldsymbol{F}_{CS}\boldsymbol{C}_S(t)\boldsymbol{F}_{VS}^T)\boldsymbol{V}_V(t) \tag{44}$$

$$\begin{aligned}\boldsymbol{I}_S(t)=&\{\dot{\boldsymbol{C}}_S(t)\boldsymbol{F}_{CS}^T\mathscr{C}^{-1}(t)+\boldsymbol{C}_S(t)\boldsymbol{F}_{CS}^T(\dot{\mathscr{C}}^{-1}(t))+\boldsymbol{C}_S(t)\boldsymbol{F}_{CS}^T\mathscr{C}^{-1}(t)\boldsymbol{A}_{CC}(t)\mathscr{C}^{-1}(t)\}\boldsymbol{q}(t)\\&+\{\boldsymbol{C}_S(t)\boldsymbol{F}_{CS}^T\mathscr{C}^{-1}(t)\boldsymbol{A}_{CL}(t)\mathscr{L}^{-1}(t)\}\boldsymbol{\phi}(t)+\{\dot{\boldsymbol{C}}_S(t)\boldsymbol{F}_{VS}^T\\&-\dot{\boldsymbol{C}}_S(t)\boldsymbol{F}_{CS}^T\mathscr{C}^{-1}(t)\boldsymbol{F}_{CS}\boldsymbol{C}_S(t)\boldsymbol{F}_{VS}^T-\boldsymbol{C}_S(t)\boldsymbol{F}_{CS}^T[(\dot{\mathscr{C}}^{-1}(t))\boldsymbol{F}_{CS}\boldsymbol{C}_S(t)\\&+\mathscr{C}^{-1}(t)\boldsymbol{F}_{CS}\dot{\boldsymbol{C}}_S(t)]\boldsymbol{F}_{VS}^T+\boldsymbol{C}_S(t)\boldsymbol{F}_{CS}^T\mathscr{C}^{-1}(t)(\boldsymbol{A}_{CV}(t)\\&-\boldsymbol{A}_{CC}(t)\mathscr{C}^{-1}(t)\boldsymbol{F}_{CS}\boldsymbol{C}_S(t)\boldsymbol{F}_{VS}^T)\}\boldsymbol{V}_V(t)+\{\boldsymbol{C}_S(t)\boldsymbol{F}_{CS}^T\mathscr{C}^{-1}(t)\boldsymbol{A}_{CI}(t)\\&-\boldsymbol{C}_S(t)\boldsymbol{F}_{CS}^T\mathscr{C}^{-1}(t)\boldsymbol{A}_{CL}(t)\mathscr{L}^{-1}(t)(\boldsymbol{F}_{\Gamma L}^T\boldsymbol{L}_{\Gamma\Gamma}(t)-\boldsymbol{L}_{L\Gamma}(t))\boldsymbol{F}_{\Gamma I}\}\boldsymbol{I}_I(t)\\&+\{\boldsymbol{C}_S(t)\boldsymbol{F}_{VS}^T-\boldsymbol{C}_S(t)\boldsymbol{F}_{CS}^T\mathscr{C}^{-1}(t)\boldsymbol{F}_{CS}\boldsymbol{C}_S(t)\boldsymbol{F}_{VS}^T\}\dot{\boldsymbol{V}}_V(t)\end{aligned} \tag{45}$$

若假定 A(i)—A(ii) 全部满足，则 $\boldsymbol{q}$，$\boldsymbol{\phi}$ 分别为电容器上的电荷及电感线圈中的磁通．在这种情况下，利用 $\boldsymbol{x}=[\boldsymbol{q}^T, \boldsymbol{\phi}^T]^T$，任何输出 $\boldsymbol{y}$ 均可用下列形式表示

$$\boldsymbol{y}(t)=\boldsymbol{C}(t)\boldsymbol{x}(t)+\boldsymbol{D}(t)\boldsymbol{u}(t).$$

(例 3)

现在来讨论图 14 所示电路，设其中各电阻、电感、电容的电阻值，电感、电容的数值均在正值范围内随时间变化，输出 $\boldsymbol{y}$ 是电容 C_3 上的电压.

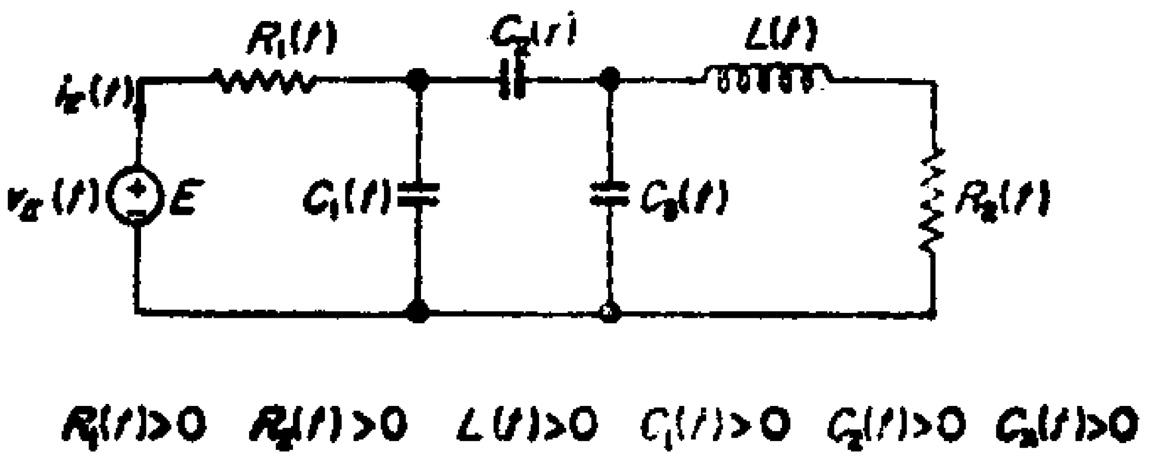


图 14　时变线性 RLC 回路

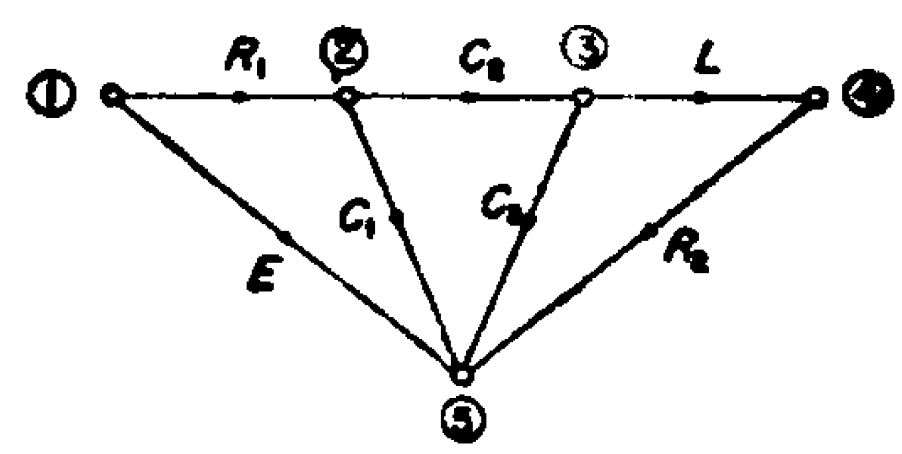


图 15　图 14 所示回路的有向图

首先，适当地规定电流的基准方向，并绘出该电路的有向图(图 15)．为了求出 N 树及与其相联系的基本割集矩阵 $\boldsymbol{Q}_f$，我们将关联矩阵 $\boldsymbol{A}_a$ 的列按着电压源枝、电容枝、电阻枝、电感枝的顺序排列

$$\boldsymbol{A}_a=\begin{array}{c}\overbrace{\begin{array}{ccccccc}E & C_1 & C_2 & C_3 & R_1 & R_2 & L\end{array}}^{\text{枝}}\\ \begin{bmatrix}1 & 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 1 & 1 & 0 & -1 & 0 & 0\\ 0 & 0 & -1 & 1 & 0 & 0 & 1\\ 0 & 0 & 0 & 0 & 0 & 1 & -1\\ -1 & -1 & 0 & -1 & 0 & -1 & 0\end{bmatrix}\end{array}\left.\begin{array}{c}① \\ ② \\ ③ \\ ④ \\ ⑤\end{array}\right\}\text{节点}$$

在 $\boldsymbol{A}_a$ 中取掉任意行，例如第 5 行，作成既约矩阵 $\boldsymbol{A}$

$$\boldsymbol{A}=\begin{array}{c}\overbrace{\begin{array}{ccccccc}E & C_1 & C_2 & C_3 & R_1 & R_2 & L\end{array}}^{\text{枝}}\\ \begin{bmatrix}1 & 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 1 & 1 & 0 & -1 & 0 & 0\\ 0 & 0 & -1 & 1 & 0 & 0 & 1\\ 0 & 0 & 0 & 0 & 0 & 1 & -1\end{bmatrix}\end{array}$$

其次，将 $\boldsymbol{A}$ 变换成标准形

$$\begin{array}{c}\overbrace{\begin{array}{ccccccc}E & C_1 & C_2 & C_3 & R_1 & R_2 & L\end{array}}^{\text{枝}}\\ \begin{bmatrix}1 & 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 1 & 0 & 1 & -1 & 0 & 1\\ 0 & 0 & 1 & -1 & 0 & 0 & -1\\ 0 & 0 & 0 & 0 & 0 & 1 & -1\end{bmatrix}\end{array}.$$

适当交换该矩阵的列，即将对应于电阻 R_2 枝的列移到第 4 列，按着表 2 变换列的顺序，得到 $\boldsymbol{Q}_f$

$$\boldsymbol{Q}_f=\begin{array}{c}\overbrace{\begin{array}{ccccccc}E & C_1 & C_2 & R_2 & C_3 & R_1 & L\end{array}}^{\text{枝}}\\ \begin{bmatrix}1 & 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 1 & 0 & 0 & 1 & -1 & 1\\ 0 & 0 & 1 & 0 & -1 & 0 & -1\\ 0 & 0 & 0 & 1 & 0 & 0 & -1\end{bmatrix}\end{array}$$

由此可见，E、C_1、C_2、R_2 的枝组成 N 树(参看图 16)．这时矩阵 $\boldsymbol{F}$ 如下

$$\boldsymbol{F}=\begin{bmatrix}\boldsymbol{F}_{VS} & \boldsymbol{F}_{VR} & \boldsymbol{F}_{VL}\\ \boldsymbol{F}_{CS} & \boldsymbol{F}_{CR} & \boldsymbol{F}_{CL}\\ \boldsymbol{0} & \boldsymbol{F}_{GR} & \boldsymbol{F}_{GL}\end{bmatrix}=\left[\begin{array}{c|c|c}0 & 1 & 0\\ \hline 1 & -1 & 1\\ -1 & 0 & -1\\ \hline 0 & 0 & -1\end{array}\right]$$

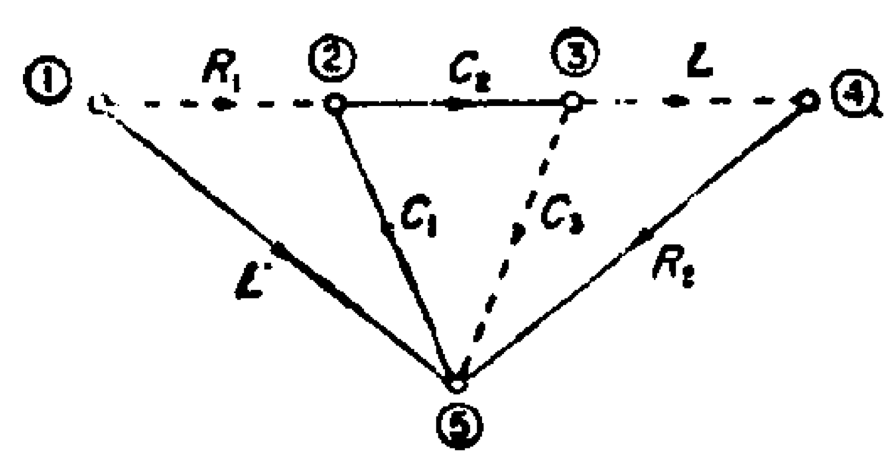


图 16　图 15 所示图的 N 树

此外，因电容、电阻、电感的电压-电流特性式分别为

$$\left[\begin{array}{c} i_{R_2}(t) \\ \hline v_{R_1}(t) \end{array}\right]=\left[\begin{array}{c|c} R_2^{-1}(t) & 0 \\ \hline 0 & R_1(t) \end{array}\right]\left[\begin{array}{c} v_{R_2}(t) \\ \hline i_{R_1}(t) \end{array}\right]$$

$$\left[\begin{array}{c} i_{C_1}(t) \\ i_{C_2}(t) \\ \hline i_{C_3}(t) \end{array}\right]=\frac{d}{dt}\left\{\left[\begin{array}{cc|c} C_1(t) & 0 & 0 \\ 0 & C_2(t) & 0 \\ \hline 0 & 0 & C_3(t) \end{array}\right]\left[\begin{array}{c} v_{C_1}(t) \\ v_{C_2}(t) \\ \hline v_{C_3}(t) \end{array}\right]\right\}$$

$$v_L(t)=\frac{d}{dt}\{L(t)i_L(t)\},$$

与(36)、(37)、(38)式对照，得

$$\boldsymbol{G}_G(t)=R_2^{-1}(t) \qquad \boldsymbol{R}_R(t)=R_1(t)$$

$$\boldsymbol{C}_C(t)=\begin{bmatrix} \mathscr{C}_1(t) & 0 \\ 0 & \mathscr{C}_2(t) \end{bmatrix} \quad \boldsymbol{C}_S(t)=C_3(t)$$

$$\boldsymbol{L}_{LL}(t)=L(t).$$

因此

$$\mathscr{C}(t)=\begin{bmatrix} C_1(t)+C_3(t) & -C_3(t) \\ -C_3(t) & C_2(t)+C_3(t) \end{bmatrix}$$

$$\mathscr{L}(t)=L(t)$$

$$\mathscr{R}(t)=R_1(t)$$

注意，$\boldsymbol{G}_G$、$\mathscr{C}$、$\mathscr{L}$、$\mathscr{R}$ 对所有的 t 均为正则．若引入(40)式定义的

$$\boldsymbol{q}(t)=\begin{bmatrix} q_1(t) \\ q_2(t) \end{bmatrix}=\begin{bmatrix} C_1(t)+C_3(t) & -C_3(t) \\ -C_3(t) & C_2(t)+C_3(t) \end{bmatrix}\begin{bmatrix} v_{C_1}(t) \\ v_{C_2}(t) \end{bmatrix}$$

$$\boldsymbol{\phi}(t)=\phi_1(t)=L(t)i_L(t),$$

根据(41)，(45)式，便得到状态方程式

$$\begin{bmatrix} \dot{q}_1(t) \\ \dot{q}_2(t) \\ \dot{\phi}_1(t) \end{bmatrix}=\begin{bmatrix} -\dfrac{C_2(t)+C_3(t)}{\Delta(t)R_1(t)} & -\dfrac{C_3(t)}{\Delta(t)R_1(t)} & -\dfrac{1}{L(t)} \\ 0 & 0 & \dfrac{1}{L(t)} \\ \dfrac{C_2(t)}{\Delta(t)} & -\dfrac{C_1(t)}{\Delta(t)} & -\dfrac{R_2(t)}{L(t)} \end{bmatrix}\begin{bmatrix} q_1(t) \\ q_2(t) \\ \phi_1(t) \end{bmatrix}$$

$$+\begin{bmatrix} \dfrac{1}{R_1(t)} \\ 0 \\ 0 \end{bmatrix}v_E(t)$$

和输出方程式

$$\boldsymbol{y}(t)=v_{C_3}(t)=\begin{bmatrix} \dfrac{C_2(t)}{\Delta(t)} & -\dfrac{C_1(t)}{\Delta(t)} & 0 \end{bmatrix}\begin{bmatrix} q_1(t) \\ q_2(t) \\ \phi_1(t) \end{bmatrix}$$

式中 $\Delta(t)\triangleq C_1(t)C_2(t)+C_1(t)C_3(t)+C_2(t)C_3(t)>0.$

一般非时变线性网络状态方程式的推导方法[38-40,42-44]

这里将要介绍一般性非时变线性网络状态方程式的推导方法．在这种网络中，除了

无源元件电阻、电容和电感以外，还包含有源元件负电阻和线性从属电源．和独立电源相对应，对于线性从属电源，可以考虑从属电压源和从属电流源．从属电压源，是其端子电压可以用网络内其它某些元件的端子电压、电流的线性组合表示的元件．从属电流源，是其端子电流可以用其它元件的端子电压、电流的线性组合表示的元件．为了使从属电源与独立电源相区别，以下用图 17(a)所示符号表示从属电压源；用图 17(b)所示符号表示从属电流源．如图 18—22 所示，利用从属电源可以将电子管、晶体管、紧耦合电感、理想变压器、回转器等价地表示出来，则这里讲的状态方程式的推导方法对包含有这些元件的最一般性的非时变线性网络均适用．

(a)从属电压源　(b)从属电流源

图 17　从属电源

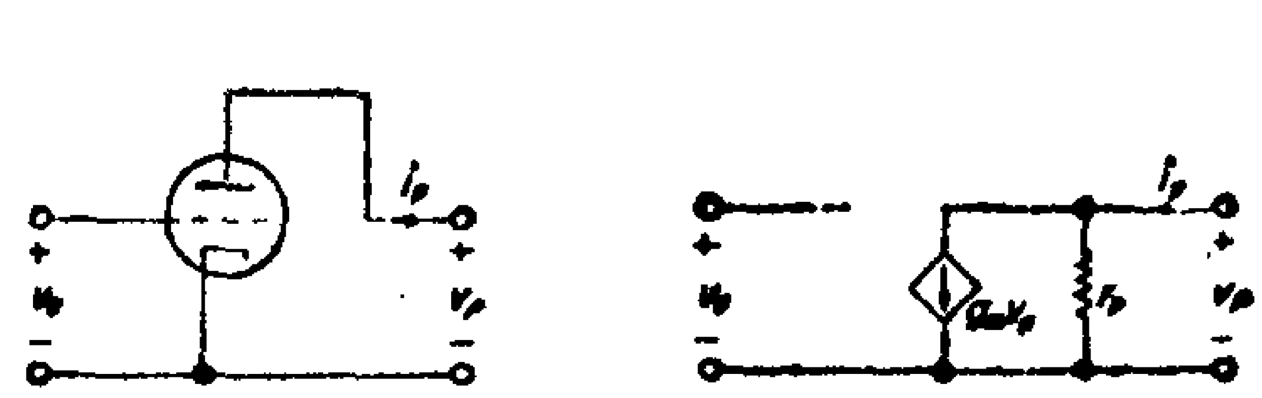


图 18　电子管小信号等效电路(例)[41]

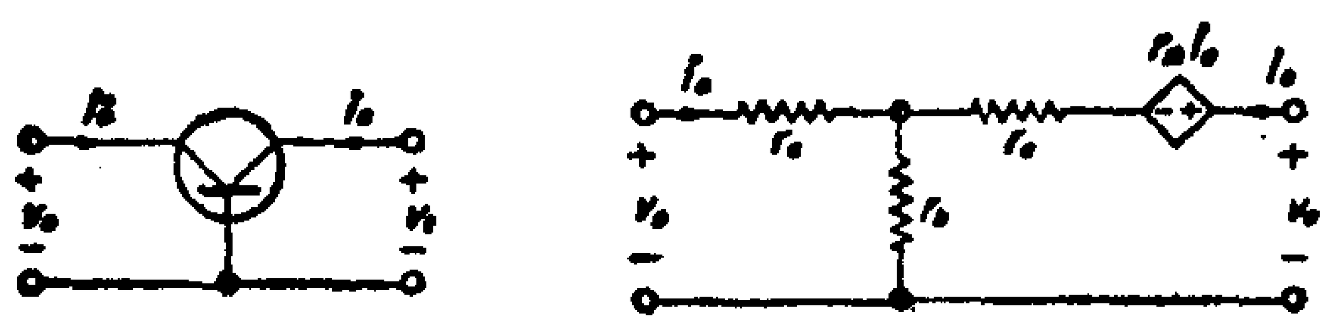


图 19　晶体管小信号等效电路(例)[41]

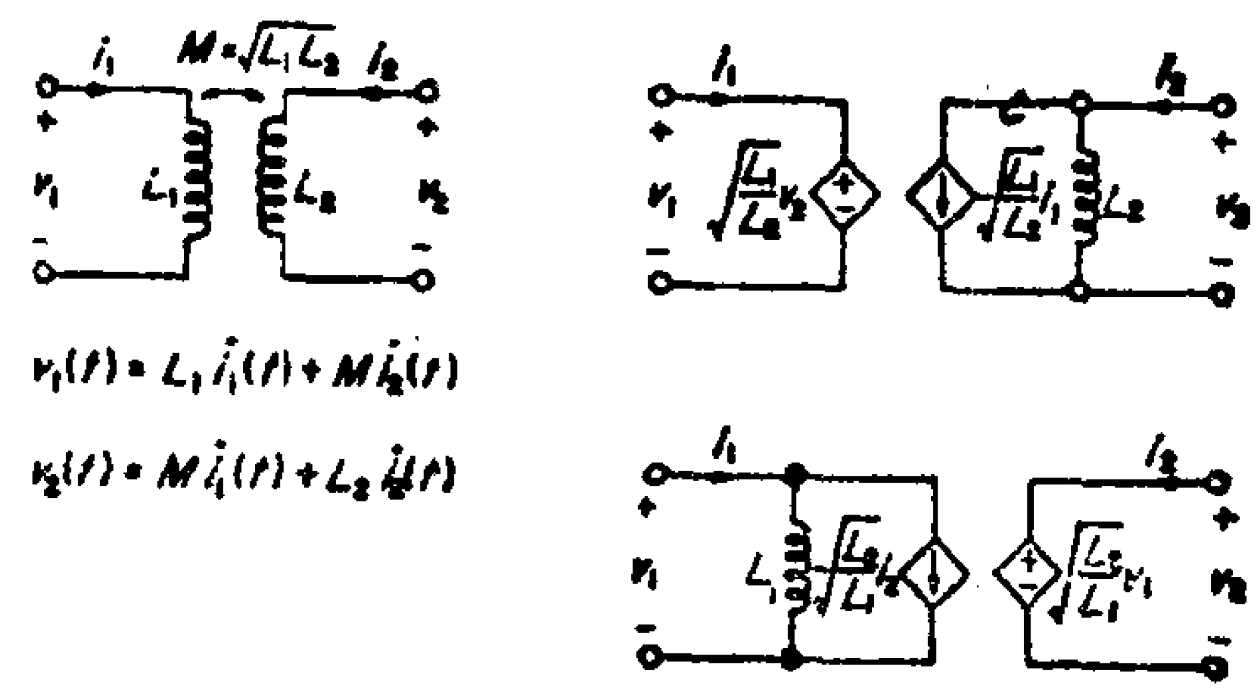


图 20　紧耦合电感等效电路

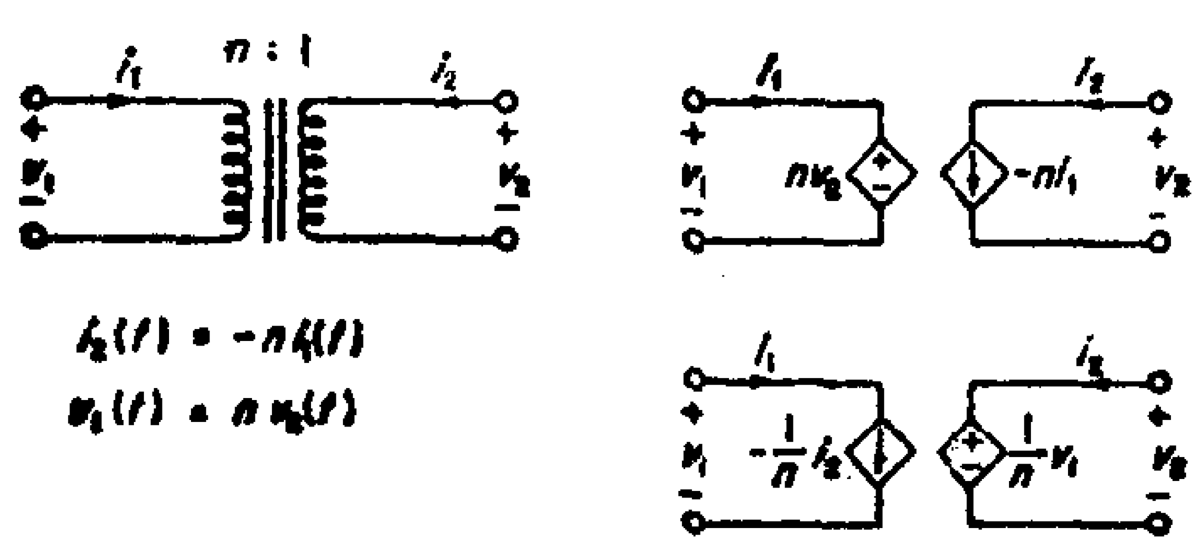


图 21　理想变压器的等效电路

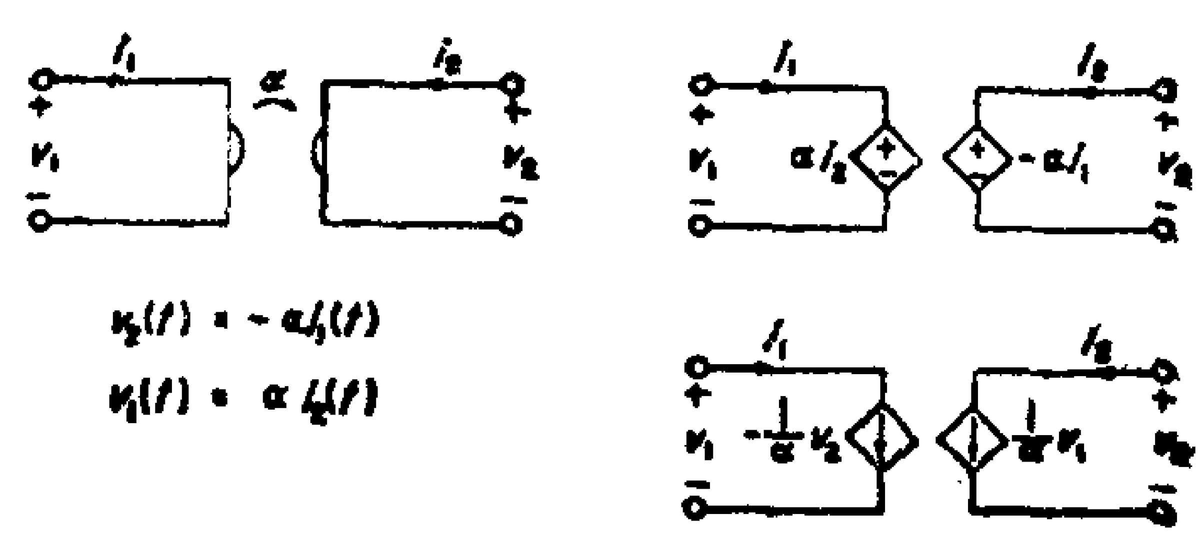


图 22　回转器的等效电路[36]

从属电源用图的枝表示时，与从属电源的端子电压或电流用其它元件的端子电压、电流的线性组合如何表示无关，和处理独立电源时一样．试回忆一下，如以前将耦合电感用图的枝表示时曾讲过，网络用图表示是为了将克希霍夫定律用相互独立的方程式的形式表示，克希霍夫定律是仅说明各元件之间连接关系的定律．因为这里完全没有考虑各元件的电压-电流特性，所以如上述这样处理从属电源是理所当然的．

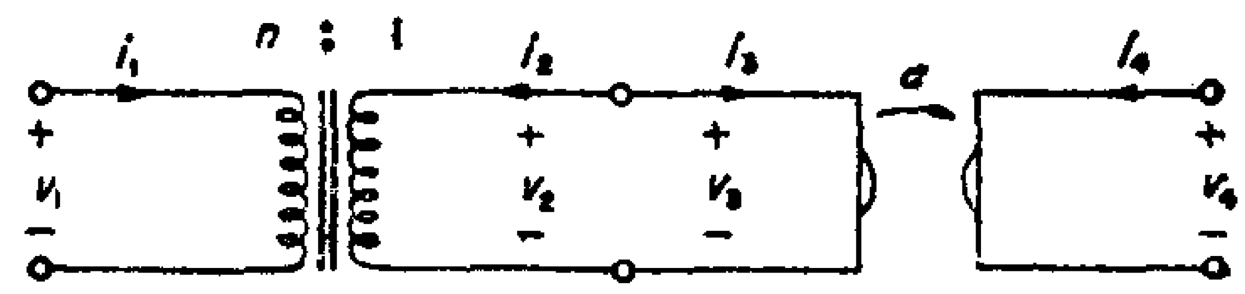
根据和以前讲过的同样理由，当然认为所讨论的网络中没有仅含有电压源的回路，但是要注意，这里除了独立的电压源以外，还有从属电压源存在，即认为下列假定成立：

A(i)　网络中没有仅含有电压源(独立或从属的电压源)的回路存在．

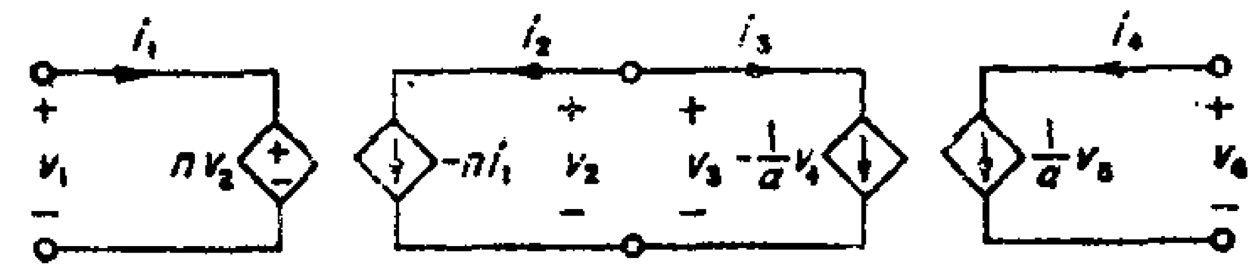
同样，关于电流源设立下列假定：

A(ii) 网络中没有仅含有电流源(独立或从属的电流源)的割集存在.

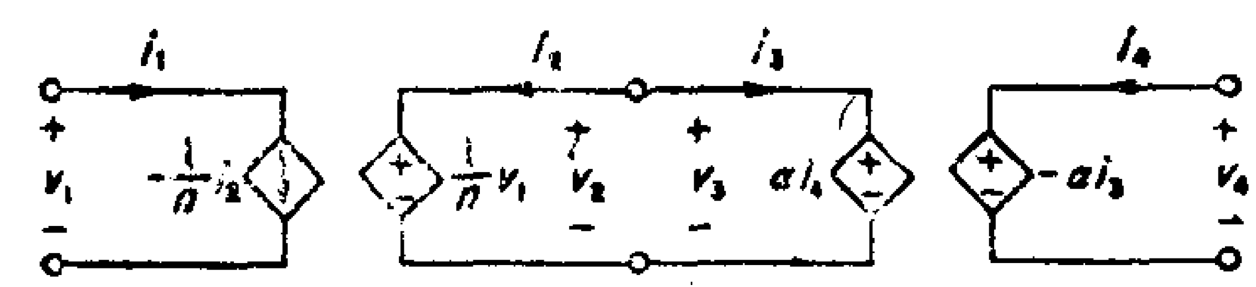
如图 21, 22 所示, 理想变压器和回转器分别均有两个等效电路, 但是必须使用满足上述假定的一个. 例如, 图 23(a)所示由理想变压器和回转器串联成的电路, 在图 23(b)的等效电路中有仅含有从属电流源的割集存在, 在图 23(c) 的等效电路中有仅含有从属电压源的回路存在, 所以图 23(b)、(c)所示等效电路不满足假定 A(i)或 A(ii). 因此, 对于图 23(a)所示电路, 必须采用图(23) (d)、(e)所示等效电路.



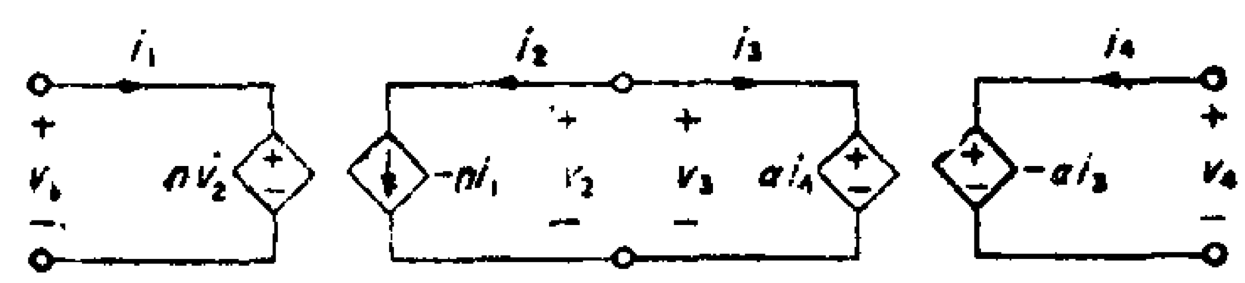
(a)理想变压器和回转器串联



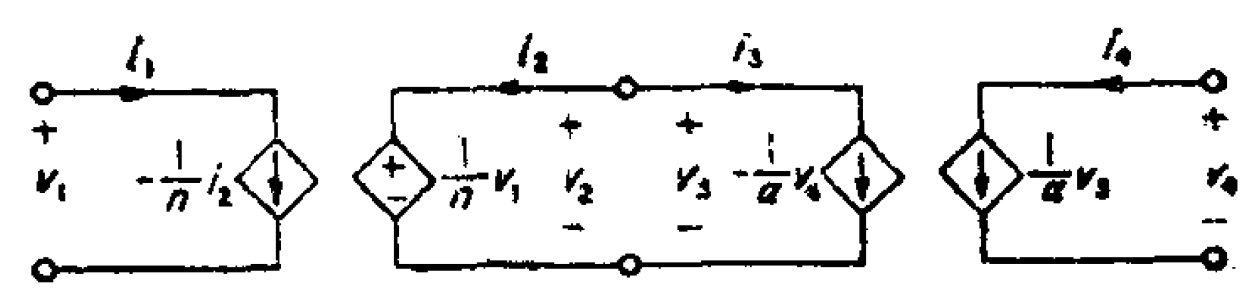
(b)(a)的不恰当的等效电路



(c) (a)的不恰当的等效电路



(d)(a)的恰当的等效电路



(e)(a)的恰当的等效电路

图 23 理想变压器和回转器串联及其等效电路

负电阻是有源元件, 但不是从属电源, 其电阻值是负的, 处理方法与具有正电阻的普通电阻完全相同.

现在, 我们来讲一下状态方程式的推导方法, 首先扩大一下 N 树的概念.

[定义 7] 包含所有独立和从属的电压源枝及尽可能多的电容枝、尽可能少的电感枝、所有独立及从属电流源枝的树, 称为规范树(normal tree, N 树).

在所给网络的有向图上, 将树选为 N 树. 按着从属电压源枝、独立电压源枝、树中电容枝、树中电阻枝、树中电感枝、树余中电容枝、树余中电阻枝、树余中电感枝、独立电流源枝、从属电流源枝的顺序对枝进行编号. 若以各枝的电压、电流为分量的向量和枝的数量用表 3 所示符号表示时, 和在讨论不一定满足假定 A(iii)—A(vi) 的 RLCM 网络时一样, $\boldsymbol{F}$ 矩阵可以分解成下列形式分块矩阵.

表3　枝的分类 III

枝的种类	电压向量	电流向量	枝的数量	
从属电压源枝	V_A	I_A	n_A	树
独立电压源枝	V_V	I_V	n_V	
树中电容枝	V_C	I_C	n_C	
树中电阻枝	V_G	I_G	n_G	
树中电感枝	V_Γ	I_Γ	n_Γ	
树余中电容枝	V_S	I_S	n_S	树余
树余中电阻枝	V_R	I_R	n_R	
树余中电感枝	V_L	I_L	n_L	
独立电流源枝	V_I	I_I	n_I	
从属电流源枝	V_B	I_B	n_B	

$$
\boldsymbol{F}=\begin{array}{c} \begin{array}{ccccc} \overbrace{\quad}^{n_S} & \overbrace{\quad}^{n_R} & \overbrace{\quad}^{n_L} & \overbrace{\quad}^{n_I} & \overbrace{\quad}^{n_B} \end{array} \\ \begin{bmatrix} \boldsymbol{F}_{AS} & \boldsymbol{F}_{AR} & \boldsymbol{F}_{AL} & \boldsymbol{F}_{AI} & \boldsymbol{F}_{AB} \\ \boldsymbol{F}_{VS} & \boldsymbol{F}_{VR} & \boldsymbol{F}_{VL} & \boldsymbol{F}_{VI} & \boldsymbol{F}_{VB} \\ \boldsymbol{F}_{CS} & \boldsymbol{F}_{CR} & \boldsymbol{F}_{CL} & \boldsymbol{F}_{CI} & \boldsymbol{F}_{CB} \\ \mathbf{0} & \boldsymbol{F}_{GR} & \boldsymbol{F}_{GL} & \boldsymbol{F}_{GI} & \boldsymbol{F}_{GB} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{F}_{\Gamma L} & \boldsymbol{F}_{\Gamma I} & \boldsymbol{F}_{\Gamma B} \end{bmatrix} \end{array} \begin{array}{l} \\ \}\, n_A \\ \}\, n_V \\ \}\, n_C \\ \}\, n_G \\ \}\, n_\Gamma, \end{array}
$$

基本割集方程式、基本回路方程式可分别表示如下

$$
\begin{bmatrix} \boldsymbol{I}_A(t) \\ \boldsymbol{I}_V(t) \\ \boldsymbol{I}_C(t) \\ \boldsymbol{I}_G(t) \\ \boldsymbol{I}_\Gamma(t) \end{bmatrix} = \begin{bmatrix} -\boldsymbol{F}_{AS} & -\boldsymbol{F}_{AR} & -\boldsymbol{F}_{AL} & -\boldsymbol{F}_{AI} & -\boldsymbol{F}_{AB} \\ -\boldsymbol{F}_{VS} & -\boldsymbol{F}_{VR} & -\boldsymbol{F}_{VL} & -\boldsymbol{F}_{VI} & -\boldsymbol{F}_{VB} \\ -\boldsymbol{F}_{CS} & -\boldsymbol{F}_{CR} & -\boldsymbol{F}_{CL} & -\boldsymbol{F}_{CI} & -\boldsymbol{F}_{CB} \\ \mathbf{0} & -\boldsymbol{F}_{GR} & -\boldsymbol{F}_{GL} & -\boldsymbol{F}_{GI} & -\boldsymbol{F}_{GB} \\ \mathbf{0} & \mathbf{0} & -\boldsymbol{F}_{\Gamma L} & -\boldsymbol{F}_{\Gamma I} & -\boldsymbol{F}_{\Gamma B} \end{bmatrix} \begin{bmatrix} \boldsymbol{I}_S(t) \\ \boldsymbol{I}_R(t) \\ \boldsymbol{I}_L(t) \\ \boldsymbol{I}_I(t) \\ \boldsymbol{I}_B(t) \end{bmatrix} \tag{46}
$$

和

$$
\begin{bmatrix} \boldsymbol{V}_S(t) \\ \boldsymbol{V}_R(t) \\ \boldsymbol{V}_L(t) \\ \boldsymbol{V}_I(t) \\ \boldsymbol{V}_B(t) \end{bmatrix} = \begin{bmatrix} \boldsymbol{F}_{AS}^T & \boldsymbol{F}_{VS}^T & \boldsymbol{F}_{CS}^T & \mathbf{0} & \mathbf{0} \\ \boldsymbol{F}_{AR}^T & \boldsymbol{F}_{VR}^T & \boldsymbol{F}_{CR}^T & \boldsymbol{F}_{GR}^T & \mathbf{0} \\ \boldsymbol{F}_{AL}^T & \boldsymbol{F}_{VL}^T & \boldsymbol{F}_{CL}^T & \boldsymbol{F}_{GL}^T & \boldsymbol{F}_{\Gamma L}^T \\ \boldsymbol{F}_{AI}^T & \boldsymbol{F}_{VI}^T & \boldsymbol{F}_{CI}^T & \boldsymbol{F}_{GI}^T & \boldsymbol{F}_{\Gamma I}^T \\ \boldsymbol{F}_{AB}^T & \boldsymbol{F}_{VB}^T & \boldsymbol{F}_{CB}^T & \boldsymbol{F}_{GB}^T & \boldsymbol{F}_{\Gamma B}^T \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_A(t) \\ \boldsymbol{V}_V(t) \\ \boldsymbol{V}_C(t) \\ \boldsymbol{V}_G(t) \\ \boldsymbol{V}_\Gamma(t) \end{bmatrix}. \tag{47}
$$

由此可见，将出现在(46)、(47)式右边的电压、电流整理后，网络中所有枝的电压、电流可以表示成下列形式(而且是代数上的)显函数

$$
\boldsymbol{W}(t) \triangleq \begin{bmatrix} \boldsymbol{V}_A(t) \\ \boldsymbol{I}_B(t) \\ \boldsymbol{V}_G(t) \\ \boldsymbol{I}_R(t) \\ \boldsymbol{I}_S(t) \\ \boldsymbol{V}_\Gamma(t) \\ \boldsymbol{V}_C(t) \\ \boldsymbol{I}_L(t) \\ \boldsymbol{V}_V(t) \\ \boldsymbol{I}_I(t) \end{bmatrix}.
$$

如前面讲过，从属电压源的电压 $\boldsymbol{V}_A$ 和从属电流源的电流 $\boldsymbol{I}_B$ 可以用网络的枝电压、电流的线性函数给出．因为所有的枝的电压、电流均为上述的 $\boldsymbol{W}$ 形线性函数，所以 $\boldsymbol{V}_A$、$\boldsymbol{I}_B$ 最后可以表示成

$$\begin{bmatrix} \boldsymbol{V}_A(t) \\ \boldsymbol{I}_B(t) \end{bmatrix} = \begin{bmatrix} \overbrace{\boldsymbol{P}_{11}}^{n_A} & \overbrace{\boldsymbol{P}_{12}}^{n_B} & \overbrace{\boldsymbol{P}_{13}}^{n_G} & \overbrace{\boldsymbol{P}_{14}}^{n_R} & \overbrace{\boldsymbol{P}_{15}}^{n_S} & \overbrace{\boldsymbol{P}_{16}}^{n_\Gamma} & \overbrace{\boldsymbol{P}_{17}}^{n_C} & \overbrace{\boldsymbol{P}_{18}}^{n_L} & \overbrace{\boldsymbol{P}_{19}}^{n_V} & \overbrace{\boldsymbol{P}_{1,10}}^{n_I} \\ \boldsymbol{P}_{21} & \boldsymbol{P}_{22} & \boldsymbol{P}_{23} & \boldsymbol{P}_{24} & \boldsymbol{P}_{25} & \boldsymbol{P}_{26} & \boldsymbol{P}_{27} & \boldsymbol{P}_{28} & \boldsymbol{P}_{29} & \boldsymbol{P}_{2,10} \end{bmatrix} \boldsymbol{W}(t) \tag{48}$$

同样，网络输出 $\boldsymbol{y}$ 也可以表示成

$$\boldsymbol{y}(t) = [\overbrace{\boldsymbol{Q}_1}^{n_A} \ \overbrace{\boldsymbol{Q}_2}^{n_B} \ \overbrace{\boldsymbol{Q}_3}^{n_G} \ \overbrace{\boldsymbol{Q}_4}^{n_R} \ \overbrace{\boldsymbol{Q}_5}^{n_S} \ \overbrace{\boldsymbol{Q}_6}^{n_\Gamma} \ \overbrace{\boldsymbol{Q}_7}^{n_C} \ \overbrace{\boldsymbol{Q}_8}^{n_L} \ \overbrace{\boldsymbol{Q}_9}^{n_V} \ \overbrace{\boldsymbol{Q}_{10}}^{n_I}] \boldsymbol{W}(t) \tag{49}$$

电阻、电容、电感的电压-电流特性可分别用下式给出

$$\begin{bmatrix} \boldsymbol{I}_G(t) \\ \boldsymbol{V}_R(t) \end{bmatrix} = \begin{bmatrix} \boldsymbol{G}_G & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{R}_R \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_G(t) \\ \boldsymbol{I}_R(t) \end{bmatrix} \tag{50}$$

$$\begin{bmatrix} \boldsymbol{I}_C(t) \\ \boldsymbol{I}_S(t) \end{bmatrix} = \begin{bmatrix} \boldsymbol{C}_C & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C}_S \end{bmatrix} \begin{bmatrix} \dot{\boldsymbol{V}}_C(t) \\ \dot{\boldsymbol{V}}_S(t) \end{bmatrix} \tag{51}$$

$$\begin{bmatrix} \boldsymbol{V}_\Gamma(t) \\ \boldsymbol{V}_L(t) \end{bmatrix} = \begin{bmatrix} \boldsymbol{L}_{\Gamma\Gamma} & \boldsymbol{L}_{\Gamma L} \\ \boldsymbol{L}_{L\Gamma} & \boldsymbol{L}_{LL} \end{bmatrix} \begin{bmatrix} \dot{\boldsymbol{I}}_\Gamma(t) \\ \dot{\boldsymbol{I}}_L(t) \end{bmatrix}. \tag{52}$$

如在简单的 RLC 网络及 RLCM 网络讲过，推导状态方程式的过程可以看成是，求电容电压、电流，电感电压、电流和独立电压源电压，独立电流源电流之间代数关系的操作过程．因此，由(46)、(47)式提出 $\boldsymbol{I}_C$、$\boldsymbol{I}_\Gamma$、$\boldsymbol{V}_S$、$\boldsymbol{V}_L$ 的关系式

$$\left.\begin{aligned} \boldsymbol{I}_C(t) &= -\boldsymbol{F}_{CS}\boldsymbol{I}_S(t) - \boldsymbol{F}_{CR}\boldsymbol{I}_R(t) - \boldsymbol{F}_{CL}\boldsymbol{I}_L(t) - \boldsymbol{F}_{CI}\boldsymbol{I}_I(t) - \boldsymbol{F}_{CB}\boldsymbol{I}_B(t) \\ \boldsymbol{I}_\Gamma(t) &= -\boldsymbol{F}_{\Gamma L}\boldsymbol{I}_L(t) - \boldsymbol{F}_{\Gamma I}\boldsymbol{I}_I(t) - \boldsymbol{F}_{\Gamma B}\boldsymbol{I}_B(t) \\ \boldsymbol{V}_S(t) &= \boldsymbol{F}_{AS}^T\boldsymbol{V}_A(t) + \boldsymbol{F}_{VS}^T\boldsymbol{V}_V(t) + \boldsymbol{F}_{CS}^T\boldsymbol{V}_C(t) \\ \boldsymbol{V}_L(t) &= \boldsymbol{F}_{AL}^T\boldsymbol{V}_A(t) + \boldsymbol{F}_{VL}^T\boldsymbol{V}_V(t) + \boldsymbol{F}_{CL}^T\boldsymbol{V}_C(t) + \boldsymbol{F}_{GL}^T\boldsymbol{V}_G(t) + \boldsymbol{F}_{\Gamma L}^T\boldsymbol{V}_\Gamma(t) \end{aligned}\right\} \tag{53}$$

由该式可见，为了推出所希望的代数关系，应该消去的变量是很明显的，即必须消去 $\boldsymbol{V}_A$、$\boldsymbol{I}_B$、$\boldsymbol{V}_G$ 和 $\boldsymbol{I}_R$．一般，消去这些变量的方法是将(46)、(47)式中 $\boldsymbol{I}_G$ 和 $\boldsymbol{V}_R$ 的关系式代入电阻的电压-电流特性式(50)，得到的

$$\left.\begin{aligned} \boldsymbol{V}_G(t) &= -\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR}\boldsymbol{I}_R(t) - \boldsymbol{G}_G^{-1}\boldsymbol{F}_{GL}\boldsymbol{I}_L(t) - \boldsymbol{G}_G^{-1}\boldsymbol{F}_{GI}\boldsymbol{I}_I(t) - \boldsymbol{G}_G^{-1}\boldsymbol{F}_{GB}\boldsymbol{I}_B(t) \\ \boldsymbol{I}_R(t) &= \boldsymbol{R}_R^{-1}\boldsymbol{F}_{GR}^T\boldsymbol{V}_G(t) + \boldsymbol{R}_R^{-1}\boldsymbol{F}_{CR}^T\boldsymbol{V}_C(t) + \boldsymbol{R}_R^{-1}\boldsymbol{F}_{VR}^T\boldsymbol{V}_V(t) + \boldsymbol{R}_R^{-1}\boldsymbol{F}_{AR}^T\boldsymbol{V}_A(t) \end{aligned}\right\} \tag{54}$$

和从属电源特性式(48)联立，对 $\boldsymbol{V}_A$、$\boldsymbol{I}_B$、$\boldsymbol{V}_G$、$\boldsymbol{I}_R$ 求解后，再代入(53)式即可．这里介绍的方法适合于采用计算机，上述操作过程可以在机器上自动地进行．

将(48)、(53)、(54)式归纳后，得

$$
\begin{bmatrix} \boldsymbol{V}_A(t) \\ \boldsymbol{I}_B(t) \\ \boldsymbol{V}_G(t) \\ \boldsymbol{I}_R(t) \\ \boldsymbol{V}_S(t) \\ \boldsymbol{I}_\Gamma(t) \\ \boldsymbol{I}_C(t) \\ \boldsymbol{V}_L(t) \end{bmatrix} = \begin{bmatrix} \boldsymbol{P}_{11} & \boldsymbol{P}_{12} & \boldsymbol{P}_{13} & \boldsymbol{P}_{14} & \boldsymbol{P}_{15} & \boldsymbol{P}_{16} & \boldsymbol{P}_{17} & \boldsymbol{P}_{18} & \boldsymbol{P}_{19} & \boldsymbol{P}_{1,10} \\ \boldsymbol{P}_{21} & \boldsymbol{P}_{22} & \boldsymbol{P}_{23} & \boldsymbol{P}_{24} & \boldsymbol{P}_{25} & \boldsymbol{P}_{26} & \boldsymbol{P}_{27} & \boldsymbol{P}_{28} & \boldsymbol{P}_{29} & \boldsymbol{P}_{2,10} \\ \boldsymbol{0} & -\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GB} & \boldsymbol{0} & -\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & -\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GL} & \boldsymbol{0} & -\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GI} \\ \boldsymbol{R}_R^{-1}\boldsymbol{F}_{AR}^T & \boldsymbol{0} & \boldsymbol{R}_R^{-1}\boldsymbol{F}_{GR}^T & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{R}_R^{-1}\boldsymbol{F}_{CR}^T & \boldsymbol{0} & \boldsymbol{R}_R^{-1}\boldsymbol{F}_{VR}^T & \boldsymbol{0} \\ \boldsymbol{F}_{AS}^T & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{F}_{CS}^T & \boldsymbol{0} & \boldsymbol{F}_{VS}^T & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{F}_{\Gamma B} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & -\boldsymbol{F}_{\Gamma L} & \boldsymbol{0} & -\boldsymbol{F}_{\Gamma I} \\ \boldsymbol{0} & -\boldsymbol{F}_{CB} & \boldsymbol{0} & -\boldsymbol{F}_{CR} & -\boldsymbol{F}_{CS} & \boldsymbol{0} & \boldsymbol{0} & -\boldsymbol{F}_{CL} & \boldsymbol{0} & -\boldsymbol{F}_{CI} \\ \boldsymbol{F}_{AL}^T & \boldsymbol{0} & \boldsymbol{F}_{GL}^T & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{F}_{\Gamma L}^T & \boldsymbol{F}_{CL}^T & \boldsymbol{0} & \boldsymbol{F}_{VL}^T & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_A(t) \\ \boldsymbol{I}_B(t) \\ \boldsymbol{V}_G(t) \\ \boldsymbol{I}_R(t) \\ \boldsymbol{I}_S(t) \\ \boldsymbol{V}_\Gamma(t) \\ \boldsymbol{V}_C(t) \\ \boldsymbol{I}_L(t) \\ \boldsymbol{V}_V(t) \\ \boldsymbol{I}_I(t) \end{bmatrix} \tag{55}
$$

由以前讲过的 RLC 网络及 RLCM 网络类推，把看成补充状态变量的电容电压 $\boldsymbol{V}_C$、$\boldsymbol{V}_S$，电感电流 $\boldsymbol{I}_L$、$\boldsymbol{I}_\Gamma$ 和输入 $\boldsymbol{V}_V$、$\boldsymbol{I}_I$ 移到右边，其余变量移到左边，再利用电容、电感的电压-电流特性式(51)、(52)消去电容电流 $\boldsymbol{I}_C$、$\boldsymbol{I}_S$，电感电压 $\boldsymbol{V}_L$、$\boldsymbol{V}_\Gamma$，便可得到

$$
\begin{bmatrix} \boldsymbol{I}-\boldsymbol{P}_{11} & -\boldsymbol{P}_{12} & -\boldsymbol{P}_{13} & -\boldsymbol{P}_{14} & -\boldsymbol{P}_{15}\boldsymbol{C}_S & -\boldsymbol{P}_{16}\boldsymbol{L}_{\Gamma\Gamma} & \boldsymbol{0} & -\boldsymbol{P}_{16}\boldsymbol{L}_{\Gamma L} \\ -\boldsymbol{P}_{21} & \boldsymbol{I}-\boldsymbol{P}_{22} & -\boldsymbol{P}_{23} & -\boldsymbol{P}_{24} & -\boldsymbol{P}_{25}\boldsymbol{C}_S & -\boldsymbol{P}_{26}\boldsymbol{L}_{\Gamma\Gamma} & \boldsymbol{0} & -\boldsymbol{P}_{26}\boldsymbol{L}_{\Gamma L} \\ \boldsymbol{0} & \boldsymbol{G}_G^{-1}\boldsymbol{F}_{GB} & \boldsymbol{I} & \boldsymbol{G}_G^{-1}\boldsymbol{F}_{GR} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ -\boldsymbol{R}_R^{-1}\boldsymbol{F}_{AR}^T & \boldsymbol{0} & -\boldsymbol{R}_R^{-1}\boldsymbol{F}_{GR}^T & \boldsymbol{I} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ -\boldsymbol{F}_{AS}^T & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{F}_{\Gamma B} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{F}_{CB} & \boldsymbol{0} & \boldsymbol{F}_{CR} & \boldsymbol{F}_{CS}\boldsymbol{C}_S & \boldsymbol{0} & \boldsymbol{C}_C & \boldsymbol{0} \\ -\boldsymbol{F}_{AL}^T & \boldsymbol{0} & -\boldsymbol{F}_{GL}^T & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{L}_{L\Gamma}-\boldsymbol{F}_{\Gamma\Gamma}^T\boldsymbol{L}_{\Gamma\Gamma} & \boldsymbol{0} & \boldsymbol{L}_{LL}-\boldsymbol{F}_{\Gamma L}^T\boldsymbol{L}_{\Gamma L} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_A(t) \\ \boldsymbol{I}_B(t) \\ \boldsymbol{V}_G(t) \\ \boldsymbol{I}_R(t) \\ \dot{\boldsymbol{V}}_S(t) \\ \dot{\boldsymbol{I}}_\Gamma(t) \\ \dot{\boldsymbol{V}}_C(t) \\ \dot{\boldsymbol{I}}_L(t) \end{bmatrix}
$$

$$
= \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{P}_{17} & \boldsymbol{P}_{18} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{P}_{27} & \boldsymbol{P}_{28} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & -\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GL} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{R}_R^{-1}\boldsymbol{F}_{CR}^T & \boldsymbol{0} \\ -\boldsymbol{I} & \boldsymbol{0} & \boldsymbol{F}_{CS}^T & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{I} & \boldsymbol{0} & -\boldsymbol{F}_{\Gamma L} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & -\boldsymbol{F}_{CL} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{F}_{CL}^T & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_S(t) \\ \boldsymbol{I}_\Gamma(t) \\ \boldsymbol{V}_C(t) \\ \boldsymbol{I}_L(t) \end{bmatrix} + \begin{bmatrix} \boldsymbol{P}_{19} & \boldsymbol{P}_{1,10} \\ \boldsymbol{P}_{29} & \boldsymbol{P}_{2,10} \\ \boldsymbol{0} & -\boldsymbol{G}_G^{-1}\boldsymbol{F}_{GI} \\ \boldsymbol{R}_R^{-1}\boldsymbol{F}_{VR}^T & \boldsymbol{0} \\ \boldsymbol{F}_{VS}^T & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{F}_{\Gamma I} \\ \boldsymbol{0} & -\boldsymbol{F}_{CI} \\ \boldsymbol{F}_{VL}^T & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_V(t) \\ \boldsymbol{I}_I(t) \end{bmatrix}. \tag{56}
$$

这是最基本的网络表达式. 若设

$$
\boldsymbol{x}_0(t) \triangleq \begin{bmatrix} \boldsymbol{V}_S(t) \\ \boldsymbol{I}_\Gamma(t) \\ \boldsymbol{V}_C(t) \\ \boldsymbol{I}_L(t) \end{bmatrix}, \quad \boldsymbol{z}(t) = \begin{bmatrix} \boldsymbol{V}_A(t) \\ \boldsymbol{I}_B(t) \\ \boldsymbol{V}_G(t) \\ \boldsymbol{I}_R(t) \end{bmatrix}, \quad \boldsymbol{u}(t) = \begin{bmatrix} \boldsymbol{V}_V(t) \\ \boldsymbol{I}_I(t) \end{bmatrix},
$$

则上式在形式上可以写成

$$
\boldsymbol{s}_0 \begin{bmatrix} \boldsymbol{z}(t) \\ \dot{\boldsymbol{x}}_0(t) \end{bmatrix} = \boldsymbol{A}_0\boldsymbol{x}_0(t) + \boldsymbol{B}_0\boldsymbol{u}(t), \tag{57}
$$

该式也可以称为中间标准形. 下面将详细讨论如何由该中间标准形推导出标准形.

由中间标准形推导标准形微分方程式的方法

当 des $\boldsymbol{S}_0 \neq 0$ 时: 若 $\boldsymbol{S}_0$ 为正则矩阵，在(57)式两边各左乘以 $\boldsymbol{S}_0^{-1}$，可立即得到标准形微分方程式

$$\dot{\boldsymbol{x}}_0(t)=\hat{\boldsymbol{A}}_0\boldsymbol{x}_0(t)+\hat{\boldsymbol{B}}_0\boldsymbol{u}(t).$$

式中 $\hat{\boldsymbol{A}}_0$ 是 $\boldsymbol{S}_0^{-1}\boldsymbol{A}_0$ 下侧的 $(n_S+n_\Gamma+n_C+n_L)\times(n_S+n_\Gamma+n_C+n_L)$ 部分矩阵，$\hat{\boldsymbol{B}}_0$ 是 $\boldsymbol{S}_0^{-1}\boldsymbol{B}_0$ 下侧的 $(n_S+n_\Gamma+n_C+n_L)\times(n_V+n_I)$ 部分矩阵.

当 $\det\boldsymbol{S}_0=0$ 时：若 $\boldsymbol{S}_0$ 不是正则矩阵，在(57)式两边各左乘以适当的行变换矩阵 $\boldsymbol{R}_0$，至少可以使 $\boldsymbol{R}_0\boldsymbol{S}_0$ 的一行(设使第 r_0 行)变为零行向量. 这时，对 $\boldsymbol{R}_0\boldsymbol{A}_0$ 和 $\boldsymbol{R}_0\boldsymbol{B}_0$ 的第 r_0 行 $\boldsymbol{a}_{r_0}$、$\boldsymbol{b}_{r_0}$ 分下列三种情况进行讨论：

1) $\boldsymbol{a}_{r_0}=\boldsymbol{0}$, $\boldsymbol{b}_{r_0}=\boldsymbol{0}$

2) $\boldsymbol{a}_{r_0}=\boldsymbol{0}$, $\boldsymbol{b}_{r_0}\neq\boldsymbol{0}$

3) $\boldsymbol{a}_{r_0}\neq\boldsymbol{0}$

情况 1)说明，在(57)式的 $(n_A+n_B+n_G+n_R+n_S+n_\Gamma+n_C+n_L)$ 个方程式中有不独立的方程式，即(57)式独立方程式的个数在 $(n_A+n_B+n_G+n_R+n_S+n_\Gamma+n_C+n_L-1)$ 以下. 但是也象 $\boldsymbol{S}_0$ 为正则矩阵的情况一样，为了得到 $\boldsymbol{x}_0$ 的标准形微分方程式，可以将(57)式对 $[\boldsymbol{z}^T, \dot{\boldsymbol{x}}_0^T]^T$ 求解. 因为 $[\boldsymbol{z}^T, \dot{\boldsymbol{x}}_0^T]^T$ 具有 $(n_A+n_B+n_G+n_R+n_S+n_\Gamma+n_C+n_L)$ 个分量，所以不能得到唯一解，$[\boldsymbol{z}^T, \dot{\boldsymbol{x}}_0^T]^T$ 的某些分量可以任意给定. 这时，因网络的特性不能唯一确定，所以我们不讨论这种情况. 图 24 是特性不能唯一确定的电路例子. 显而易见，该电路的特性不是唯一的，而且当 v_d、i_d、v_R、i_R 中任意一个给定之后，其它变量才能唯一确定.

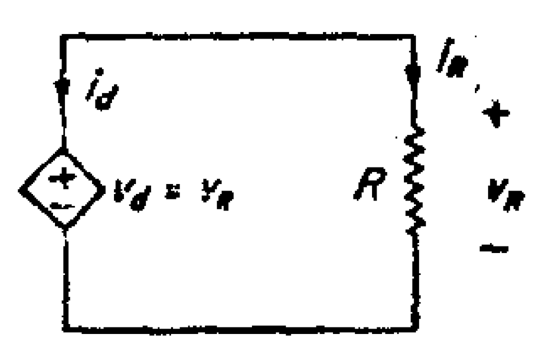


图 24 特性不能唯一确定的电路

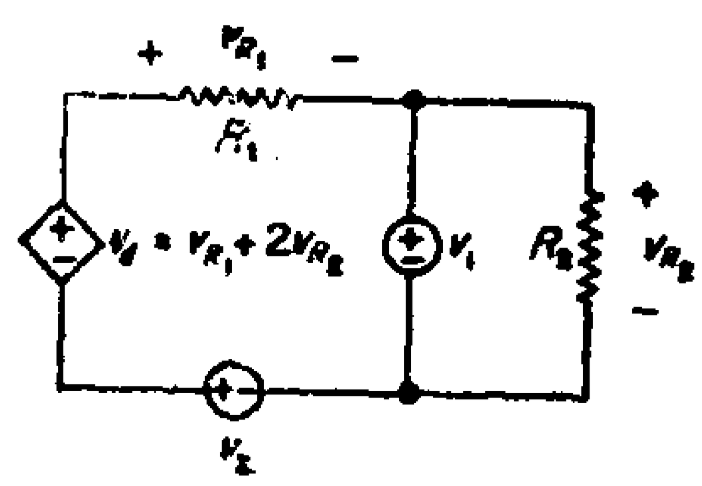


图 25 独立电源之间发生从属关系的电路

情况 2)说明，输入，即独立电压源的电压和独立电流源的电流之间存在着从属关系. 因为这与电源的独立性相矛盾，所以不讨论这种情况. 最初曾假定，没有仅含有独立电压源的回路或仅含有独立电流源的割集. 这只不过是根据网络的拓扑学结构假定独立电源之间不发生从属关系. 但是在我们现在所讨论的包含有从属电源的网络中，根据元件的结合方式，在独立电源之间可能产生从属关系. 图 25 所示电路是其中的一个例子. 利用克希霍夫第二定律列出该电路的回路方程式，得

$$v_d=v_{R_1}+v_1-v_2$$

但因从属电压源的电压 v_d 由下式决定

$$v_d=v_{R_1}+2v_{R_2}=v_{R_1}+2v_1$$

则必须

$$v_2=-v_1$$

这与 v_1 和 v_2 的独立性相矛盾.

情况 3)，在(57)式两边各左乘以行变换矩阵 $\boldsymbol{R}_0$，得

$$\boldsymbol{R}_0\boldsymbol{S}_0\begin{bmatrix}\boldsymbol{z}(t)\\ \dot{\boldsymbol{x}}_0(t)\end{bmatrix}=\boldsymbol{R}_0\boldsymbol{A}_0\boldsymbol{x}_0(t)+\boldsymbol{R}_0\boldsymbol{B}_0\boldsymbol{u}(t) \tag{58}$$

其第 r_0 式

$$0=\boldsymbol{a}_{r_0}\boldsymbol{x}_0(t)+\boldsymbol{b}_{r_0}\boldsymbol{u}(t) \tag{59}$$

表明，$\boldsymbol{x}_0$ 至少有一个分量可以由 $\boldsymbol{x}_0$ 的其它分量和输入 $\boldsymbol{u}$ 的线性组合表示. 将(59)式对

$\boldsymbol{x}_0$ 的该分量解出，代入(58)式中除第 r_0 式以外的其它式子. 与(57)式比较，变量及方程式的个数都要少一个，可以得到下列形式的网络表示式

$$\boldsymbol{S}_1\begin{bmatrix}\boldsymbol{z}(t)\\ \dot{\boldsymbol{x}}_1(t)\end{bmatrix}=\boldsymbol{A}_1\boldsymbol{x}_1(t)+\boldsymbol{B}_1\boldsymbol{u}(t)+\boldsymbol{B}_1^{(1)}\dot{\boldsymbol{u}}(t). \tag{60}$$

如上式所示，在这里网络的表示中有可能出现输入的微分. 显然，若 $\boldsymbol{S}_1$ 为正则矩阵，则和讨论(57)式中的 $\boldsymbol{S}_0$ 一样，在(60)式两边各左乘以 $\boldsymbol{S}_1^{-1}$，可以得到包含输入微分的标准形微分方程式. 另外，若 $\boldsymbol{S}_1$ 不是正则矩阵，则也和讨论(57)式时一样，在(60)式两边各左乘以适当的行变换矩阵 $\boldsymbol{R}_1$，使 $\boldsymbol{R}_1\boldsymbol{S}_1$ 的某一行（设为第 r_1 行）变为零行向量. 下面我们对 $\boldsymbol{R}_1\boldsymbol{A}_1$、$\boldsymbol{R}_1\boldsymbol{B}_1$、$\boldsymbol{R}_1\boldsymbol{B}^{(1)}$ 的第 r_1 行 $\boldsymbol{a}_{r_1}$、$\boldsymbol{b}_{r_1}$、$\boldsymbol{b}_{r_1}^{(1)}$ 分下列三种情况进行讨论

1)′ $\boldsymbol{a}_{r_1}=\boldsymbol{0}$，$\boldsymbol{b}_{r_1}=\boldsymbol{0}$，$\boldsymbol{b}_{r_1}^{(1)}=\boldsymbol{0}$

2)′ $\boldsymbol{a}_{r_1}=\boldsymbol{0}$，$\boldsymbol{b}_{r_1}\neq\boldsymbol{0}$ 或 $\boldsymbol{b}_{r_1}^{(1)}\neq\boldsymbol{0}$

3)′ $\boldsymbol{a}_{r_1}\neq\boldsymbol{0}$.

情况 1)′、2)′ 不能和(57)式的情况 1)、2)同样地进行讨论，而情况 3)′ 可很容易地由前述类推，变量个数及方程式个数均比(60)式少一个，可以得到下列形式的网络表达式

$$\boldsymbol{S}_2\begin{bmatrix}\boldsymbol{z}(t)\\ \dot{\boldsymbol{x}}_2(t)\end{bmatrix}=\boldsymbol{A}_2\boldsymbol{x}_2(t)+\boldsymbol{B}_2\boldsymbol{u}(t)+\boldsymbol{B}_2^{(1)}\dot{\boldsymbol{u}}(t)+\boldsymbol{B}_2^{(2)}\boldsymbol{u}^{(2)}(t)$$

该过程可以反复进行 n 次($n\leqslant n_S+n_\Gamma+n_C+n_L$)，最后得到下列网络表达式

$$\boldsymbol{S}_n\begin{bmatrix}\boldsymbol{z}(t)\\ \dot{\boldsymbol{x}}_n(t)\end{bmatrix}=\boldsymbol{A}_n\boldsymbol{x}_n+\boldsymbol{B}_n\boldsymbol{u}(t)+\boldsymbol{B}_n^{(1)}\dot{\boldsymbol{u}}(t)+\cdots+\boldsymbol{B}_n^{(n)}\boldsymbol{u}^{(n)}(t). \tag{61}$$

若式中 $\boldsymbol{S}_n$ 为正则矩阵，由(61)式显而易见，可以得到下列包含高阶输入微分的标准形微分方程式

$$\dot{\boldsymbol{x}}_n(t)=\hat{\boldsymbol{A}}_n\boldsymbol{x}_n(t)+\hat{\boldsymbol{B}}_n\boldsymbol{u}(t)+\hat{\boldsymbol{B}}_n^{(1)}\dot{\boldsymbol{u}}(t)+\cdots+\hat{\boldsymbol{B}}_n^{(n)}\boldsymbol{u}_n^{(n)}(t) \tag{62}$$

式中 $\boldsymbol{A}_n$、$\hat{\boldsymbol{B}}_n$、$\hat{\boldsymbol{B}}_n^{(1)}$、…、$\hat{\boldsymbol{B}}_n^{(n)}$ 分别为由 $\boldsymbol{S}_n^{-1}\boldsymbol{A}_n$、$\boldsymbol{S}_n^{-1}\boldsymbol{B}_n$、$\boldsymbol{S}_n^{-1}\boldsymbol{B}_n^{(1)}$、…、$\boldsymbol{S}_n^{-1}\boldsymbol{B}_n^{(n)}$ 下侧($n_S+n_\Gamma+n_C+n_L-n$)行组成的矩阵.

其次我们来看，如何利用(62)式的变量 $\boldsymbol{x}_n$ 和输入 $\boldsymbol{u}$ 表示(49)式确定的输出 $\boldsymbol{y}$.

为此，先看 $\boldsymbol{W}$ 如何用 $\boldsymbol{x}_n$ 和 $\boldsymbol{u}$ 表示. 因 $\boldsymbol{W}$ 中 $\boldsymbol{V}_A$、$\boldsymbol{I}_B$、$\boldsymbol{V}_G$、$\boldsymbol{I}_R$ 包含在 $\boldsymbol{z}$ 里，由(61)式可见，它可以用 $\boldsymbol{x}_n$、$\boldsymbol{u}$ 及 $\boldsymbol{u}$ 的高阶微分表示. 根据以前的定义，$\boldsymbol{V}_C$ 和 $\boldsymbol{I}_L$ 包含在 $\boldsymbol{x}_0$ 中. 因 $\boldsymbol{x}_0$ 中除去用其它分量和输入(以及输入的高阶微分)表示的分量后是 $\boldsymbol{x}_n$，所以 $\boldsymbol{x}_0$ 应由 $\boldsymbol{x}_n$、$\boldsymbol{u}$ 及 $\boldsymbol{u}$ 的高阶微分表示，$\boldsymbol{V}_C$、$\boldsymbol{I}_L$ 也应由 $\boldsymbol{x}_n$、$\boldsymbol{u}$ 及 $\boldsymbol{u}$ 的高阶微分表示. 由电压-电流特性式(50)～(52)、基本割集方程式(46)、基本回路方程式(47)可见，$\boldsymbol{I}_S$ 和 $\boldsymbol{V}_\Gamma$ 可以表示成

$$\begin{aligned}\boldsymbol{I}_S(t)&=\boldsymbol{C}_S\dot{\boldsymbol{V}}_S(t)=\boldsymbol{C}_S\{\boldsymbol{F}_{AS}^T\dot{\boldsymbol{V}}_A(t)+\boldsymbol{F}_{VS}^T\dot{\boldsymbol{V}}_V(t)+\boldsymbol{F}_{CS}^T\dot{\boldsymbol{V}}_C(t)\}\\ \boldsymbol{V}_\Gamma(t)&=\boldsymbol{L}_{\Gamma\Gamma}\dot{\boldsymbol{I}}_\Gamma(t)+\boldsymbol{L}_{\Gamma L}\dot{\boldsymbol{I}}_L(t)\\ &=\boldsymbol{L}_{\Gamma\Gamma}\{-\boldsymbol{F}_{\Gamma L}\dot{\boldsymbol{I}}_L(t)-\boldsymbol{F}_{\Gamma I}\dot{\boldsymbol{I}}_I(t)-\boldsymbol{F}_{\Gamma B}\dot{\boldsymbol{I}}_B(t)\}+\boldsymbol{L}_{\Gamma L}\dot{\boldsymbol{I}}_L(t).\end{aligned}$$

如上所述，因 $\boldsymbol{V}_A$、$\boldsymbol{I}_B$、$\boldsymbol{V}_C$、$\boldsymbol{I}_L$ 可由 $\boldsymbol{x}_n$、$\boldsymbol{u}$ 及 $\boldsymbol{u}$ 的高阶微分表示，则 $\dot{\boldsymbol{V}}_A$、$\dot{\boldsymbol{I}}_B$、$\dot{\boldsymbol{V}}_C$、$\dot{\boldsymbol{I}}_L$，因而 $\boldsymbol{I}_S$、$\boldsymbol{V}_\Gamma$ 也可以用 $\boldsymbol{x}_n$、$\boldsymbol{u}$ 及 $\boldsymbol{u}$ 的高阶微分表示. 由以上可见，一般输出 $\boldsymbol{y}$ 可以表示成下列形式

$$\boldsymbol{y}(t)=\boldsymbol{C}\boldsymbol{x}_n(t)+\boldsymbol{D}\boldsymbol{u}(t)+\boldsymbol{D}^{(1)}\dot{\boldsymbol{u}}(t)+\cdots+\boldsymbol{D}^{(m)}\boldsymbol{u}^{(m)}(t). \tag{63}$$

因此，若将 $\boldsymbol{u}$ 的(高阶)微分也作为输入，则 $\boldsymbol{x}_n$ 满足状态变量的条件，可以说(62)式是状态方程式.

现在我们举个简单例子，来说明如何用以上方法由中间标准形求标准形微分方程式.

(例 4) 试看下列中间标准形

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} z(t) \\ \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix}=\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 2 & 0 \end{bmatrix}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}+\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}u(t).$$

将该式两边各左乘以使左边系数矩阵变成行标准形的变换矩阵

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix}=\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix},$$

得

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} z(t) \\ \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix}=\begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}+\begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}u(t)$$

因此左边不是正则系数矩阵. 因第 3 行 $[x_1,\ x_2]^T$ 的系数向量不是零行向量，若将第 3 行取出改写成

$$2x_1(t)+x_2(t)+2u(t)=0,$$

解出 $x_2(t)$，得

$$x_2(t)=-2x_1(t)-2u(t).$$

将其代入第 2 行整理(因第 1 行中不出现 x_2、$\dot{x}_2$，则第 1 行原样不动)后，得

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} z(t) \\ \dot{x}_1(t) \end{bmatrix}=\begin{bmatrix} 0 \\ 1 \end{bmatrix}x_1(t)+\begin{bmatrix} -1 \\ 1 \end{bmatrix}u(t)+\begin{bmatrix} 0 \\ 2 \end{bmatrix}\dot{u}(t).$$

因该式左边是正则系数矩阵，则其可以变换成单位矩阵

$$\begin{bmatrix} z(t) \\ \dot{x}_1(t) \end{bmatrix}=\begin{bmatrix} 0 \\ -1 \end{bmatrix}x_1(t)+\begin{bmatrix} -1 \\ -1 \end{bmatrix}u(t)+\begin{bmatrix} 0 \\ -2 \end{bmatrix}\dot{u}(t)$$

第 2 行是包含输入微分的标准形微分方程式

$$\dot{x}_1(t)=-x_1(t)-u(t)-2\dot{u}(t).$$

但是，用上述方法判断能否由中间标准形求得标准形状态方程式，实际上只是消去非独立变量的过程. 在不能求得标准形的情况下，将会消耗大量的计算工作. 为了避免这种无益的计算，希望能够通过观察所给出的中间标准形直接判断能否得到标准形. 因此，下面将给出根据中间标准形的系数矩阵进行这种判断的定理. 证明省略.

[定理 3][42~44] 将中间标准形(57)式左边的系数矩阵 $\boldsymbol{S}_0$ 分解成左边的 $(n_A+n_B+n_G+n_R)$ 列和右边余下的 $(n_S+n_T+n_C+n_L)$ 列两个部分矩阵，表示成

$$\boldsymbol{S}_0=[\boldsymbol{S}_{01}:\boldsymbol{S}_{02}].$$

这时，由中间标准形能够得到标准形状态方程式(也包含输入微分)的充分与必要条件是，$\boldsymbol{S}$ 多项式

$$\boldsymbol{D}(s)\triangleq\det[\boldsymbol{S}_{01}:s\boldsymbol{S}_{02}-\boldsymbol{A}_0]$$

不恒等于零(即 $D(s)$ 不是多项式环 $P(s,\ R)$ 的零元). 而且,若能求得标准形状态方程式,则状态向量的维数和 $D(s)$ 的次数一致.

最后谈一下,如何由包含有高阶输入微分的标准形微分方程式(62)推导出不含有输入微分的标准形微分方程式,后者在数值计算等方面较为方便. 若引入新的变量

$$\boldsymbol{x}_{n+1}(t)=\boldsymbol{x}_n(t)-\hat{\boldsymbol{B}}_n^{(n)}\boldsymbol{u}^{(n-1)}(t),$$

因

$$\dot{\boldsymbol{x}}_{n+1}(t)=\dot{\boldsymbol{x}}_n(t)-\hat{\boldsymbol{B}}_n^{(n)}\boldsymbol{u}^{(n)}(t)$$

则 $\boldsymbol{x}_{n+1}$ 满足下式

$$\dot{\boldsymbol{x}}_{n+1}(t)=\hat{\boldsymbol{A}}_n\boldsymbol{x}_{n+1}(t)+\hat{\boldsymbol{B}}_n\boldsymbol{u}(t)+\hat{\boldsymbol{B}}_n^{(1)}\dot{\boldsymbol{u}}(t)+\cdots+\{\hat{\boldsymbol{B}}_n^{(n-1)}+\hat{\boldsymbol{A}}_n\hat{\boldsymbol{B}}_n^{(n)}\}\boldsymbol{u}^{(n-1)}(t).$$

注意,(62)式用输入的 n 阶微分表示,而该微分方程式用输入的 $(n-1)$ 阶微分表示. 其次,若引入新的变量

$$\boldsymbol{x}_{n+2}(t)=\boldsymbol{x}_{n+1}(t)-\{\hat{\boldsymbol{B}}_n^{(n-1)}+\hat{\boldsymbol{A}}_n\hat{\boldsymbol{B}}_n^{(n)}\}\boldsymbol{u}^{(n-2)}(t)$$

显然最高只出现输入的 $(n-2)$ 阶微分,可以得到 $\boldsymbol{x}_{n+2}$ 的微分方程式. 反复进行该过程,最终得到下列标准形微分方程式

$$\dot{\boldsymbol{x}}_{2n}(t)=\hat{\boldsymbol{A}}_n\boldsymbol{x}_{2n}(t)+\{\hat{\boldsymbol{B}}_n+\hat{\boldsymbol{A}}_n\{\hat{\boldsymbol{B}}_n^{(1)}+\hat{\boldsymbol{A}}_n\{\hat{\boldsymbol{B}}_n^{(2)}+\hat{\boldsymbol{A}}_n\{\cdots+\hat{\boldsymbol{A}}_n\hat{\boldsymbol{B}}_n^{(n)}\}\cdots\}\boldsymbol{u}(t). \tag{64}$$

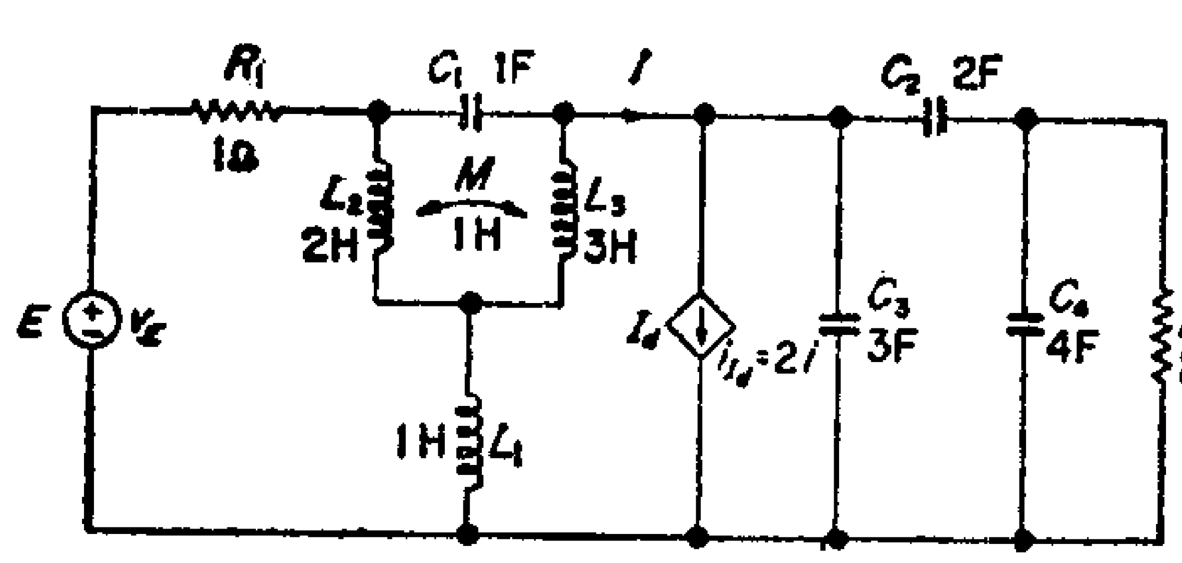
对于新的变量 $\boldsymbol{x}_{2n}$, 输出 $\boldsymbol{y}$ 可以表示成

$$\boldsymbol{y}(t)=\begin{cases}\boldsymbol{C}\boldsymbol{x}_{2n}(t)+\cdots+\{\boldsymbol{D}^{(n-2)}+\boldsymbol{C}\{\hat{\boldsymbol{B}}_n^{(n-1)}+\hat{\boldsymbol{A}}_n\hat{\boldsymbol{B}}_n^{(n)}\}\}\boldsymbol{u}^{(n-2)}(t)+\{\boldsymbol{D}^{(n-1)}\\ \quad +\boldsymbol{C}\hat{\boldsymbol{B}}_n^{(n)}\}\boldsymbol{u}^{(n-1)}(t)+\boldsymbol{D}^{(n)}\boldsymbol{u}^{(n)}(t)+\cdots+\boldsymbol{D}^{(m)}\boldsymbol{u}^{(m)}(t), \qquad m\geqslant n\\ \boldsymbol{C}\boldsymbol{x}_{2n}(t)+\cdots+\{\boldsymbol{D}^{(m)}+\boldsymbol{C}\{\hat{\boldsymbol{B}}_n^{(m+1)}+\hat{\boldsymbol{A}}_n\{\boldsymbol{B}_n^{(m+2)}+\hat{\boldsymbol{A}}_n\{\cdots+\hat{\boldsymbol{A}}_n\hat{\boldsymbol{B}}_n^{(n)}\}\cdots\}\boldsymbol{u}^{(m)}(t)\\ \quad +\boldsymbol{C}\{\boldsymbol{B}_n^{(m+2)}+\hat{\boldsymbol{A}}_n\{\hat{\boldsymbol{B}}_n^{(m+3)}+\hat{\boldsymbol{A}}_n\{\cdots+\hat{\boldsymbol{A}}_n\hat{\boldsymbol{B}}_n^{(n)}\}\cdots\}\boldsymbol{u}^{(m+1)}(t)+\cdots\\ \quad +\boldsymbol{C}\hat{\boldsymbol{B}}_n^{(n)}\boldsymbol{u}^{(n-1)}(t). \qquad m<n\end{cases} \tag{65}$$

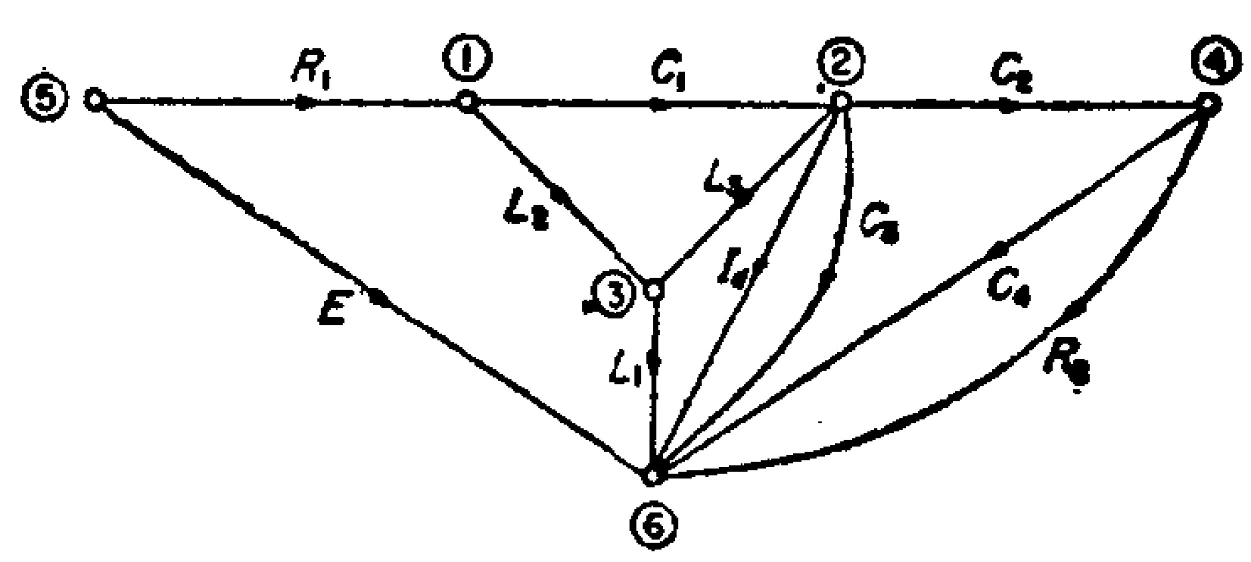
若 $\boldsymbol{u}$ 的(高阶)微分也视为输入,则和 $\boldsymbol{x}_n$ 一样, $\boldsymbol{x}_{2n}$ 也满足状态变量的条件,(64)式是状态方程式. $\boldsymbol{x}_n$ 以电容电压及电感电流为其分量,而 $\boldsymbol{x}_{2n}$ 可能已经不是这样具有明显物理意义的变量.

现在我们举个具体例子,说明如何推导上述状态方程式.

(例5) 试看图26(a)所示电路,设从属电流源的电流为输出. 同图(b)为对应于该电路的有向图. 首先,从关联矩阵中取掉对应于节点⑥的行,写出既约关联矩阵 $\boldsymbol{A}$. 这时,为了按定理3的步骤找出 N 树,将 $\boldsymbol{A}$ 的列按着对应于独立电压源枝、电容枝、电阻枝、电感枝、从属电流源枝的顺序排列,得到



(a)包含从属电源的非时变线性电路



(b)对应于(a)上电路的有向图

图26 包含从属电源的非时变线性电路及其对应的有向图

$$
\boldsymbol{A}=\begin{array}{c}
\overbrace{\begin{array}{ccccccccccc} E & C_1 & C_2 & C_3 & C_4 & R_1 & R_2 & L_1 & L_2 & L_3 & I_d \end{array}}^{\text{枝}} \\
\begin{bmatrix}
0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\
0 & -1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 & 0 \\
0 & 0 & -1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\end{array}.
$$

为了利用基本行变换及列变换将 $\boldsymbol{A}$ 变换成对应于 N 树的基本割集矩阵 $\boldsymbol{Q}_f \triangleq [\boldsymbol{I}:\boldsymbol{F}]$，首先用基本行变换将 $\boldsymbol{A}$ 变换成行标准形

$$
\boldsymbol{A}\text{ 的行标准形}=\begin{array}{c}
\overbrace{\begin{array}{ccccccccccc} E & C_1 & C_2 & C_3 & C_4 & R_1 & R_2 & L_1 & L_2 & L_3 & I_d \end{array}}^{\text{枝}} \\
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & -1 & 0 & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & -1 & 1 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 & 0
\end{bmatrix}
\end{array}.
$$

为了在 $\boldsymbol{A}$ 的标准形左边组成 5×5 阶单位矩阵，将对应于 L_1 的列移到第 5 列即可[1)]. 为此，将对应于 L_1 的列依次与对应于 R_2、R_1、C_4 的列进行交换，便得到如下基本割集矩阵

$$
\boldsymbol{Q}_f=\begin{array}{c}
\overbrace{\begin{array}{ccccc:cccccc} E & C_1 & C_2 & C_3 & L_1 & C_4 & R_1 & R_2 & L_2 & L_3 & I_d \end{array}}^{\text{枝}} \\
\left[\begin{array}{ccccc:cccccc}
1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & -1 & 0 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 1 & -1 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 & -1 & 0
\end{array}\right]
\end{array}.
$$

$$\triangleq [\boldsymbol{I}:\boldsymbol{F}]$$

由此可见，N 树由 E、C_1、C_2、C_3、L_1 枝组成(参看图 27)。

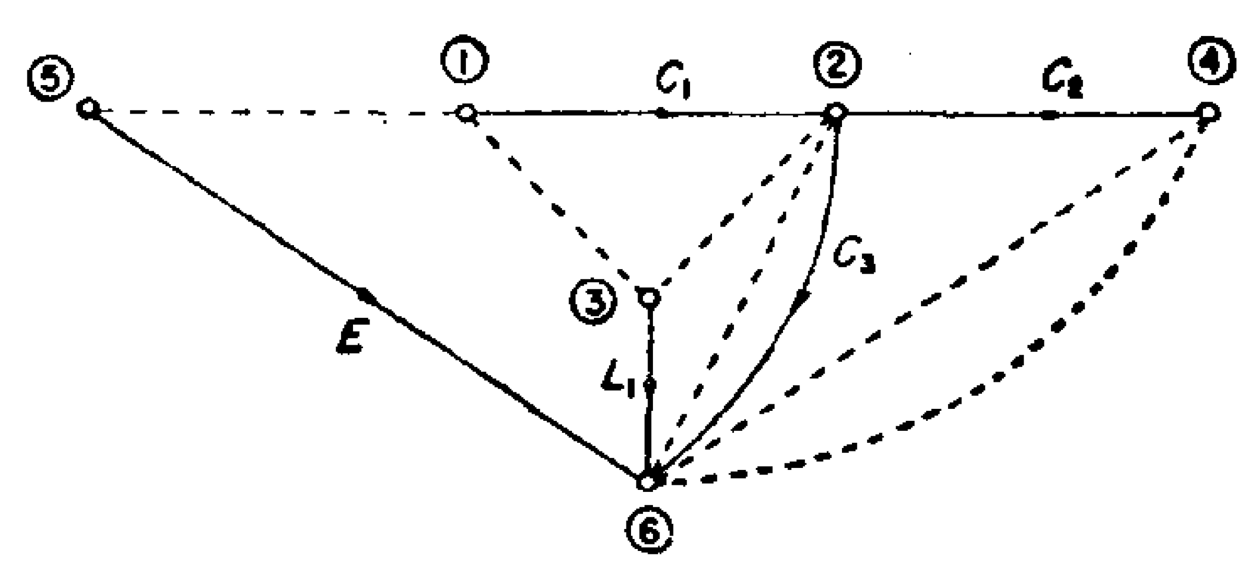


图 27　图 26(b)所示图的 N 树

$\boldsymbol{F}$ 可以分解成如下分块矩阵

1) 由 $\boldsymbol{A}$ 的行标准形可见，与 E、C_1、C_2、I_d、L_1 对应的列也可以组成 5×5 阶单位矩阵，即 E_1、C_1、C_2、I_d、L_1 的枝构成树. 但是该树包含从属电流源枝，所以不是 N 树.

$$
\boldsymbol{F}=\begin{bmatrix} \boldsymbol{F}_{VS} & \boldsymbol{F}_{VR} & \boldsymbol{F}_{VL} & \boldsymbol{F}_{VB} \\ \boldsymbol{F}_{CS} & \boldsymbol{F}_{CR} & \boldsymbol{F}_{CL} & \boldsymbol{F}_{CB} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{F}_{\Gamma L} & \boldsymbol{F}_{\Gamma B} \end{bmatrix}=\left[\begin{array}{c:cc:cc:c} 0 & 1 & 0 & 0 & 0 & 0 \\ \hdashline 0 & -1 & 0 & 1 & 0 & 0 \\ -1 & 0 & -1 & 0 & 0 & 0 \\ 1 & -1 & 1 & 1 & 1 & 1 \\ \hdashline 0 & 0 & 0 & -1 & -1 & 0 \end{array}\right]. \tag{66}
$$

基本割集方程式可以写成下列形式

$$
\begin{bmatrix} i_E(t) \\ i_{C_1}(t) \\ i_{C_2}(t) \\ i_{C_3}(t) \\ i_{L_1}(t) \end{bmatrix}=\begin{bmatrix} 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ -1 & 1 & -1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}\begin{bmatrix} i_{C_4}(t) \\ i_{R_1}(t) \\ i_{R_2}(t) \\ i_{L_2}(t) \\ i_{L_3}(t) \\ i_{I_d}(t) \end{bmatrix}.
$$

因此，从属电流源的电流 i_{I_d} 可以写成

$$
i_{I_d}(t)=2i(t)\triangleq 2\{i_{C_1}(t)-i_{L_3}(t)\}=2\{i_{R_1}(t)-i_{L_2}(t)-i_{L_3}(t)\}
$$

$$
=[0\,\vdots\,2\;0\,\vdots\,0\,\vdots\,0\,\vdots\,0\;0\;0\,\vdots\,-2\;\;-2\,\vdots\,0]\left[\begin{array}{c} i_{I_d}(t) \\ \hdashline i_{R_1}(t) \\ i_{R_2}(t) \\ \hdashline i_{C_4}(t) \\ \hdashline v_{L_1}(t) \\ \hdashline v_{C_1}(t) \\ v_{C_2}(t) \\ v_{C_3}(t) \\ \hdashline i_{L_2}(t) \\ i_{L_3}(t) \\ \hdashline v_E(t) \end{array}\right]
$$

将该式与(48)式比较，得

$$
[\boldsymbol{P}_{22}\,\vdots\,\boldsymbol{P}_{24}\,\vdots\,\boldsymbol{P}_{25}\,\vdots\,\boldsymbol{P}_{26}\,\vdots\,\boldsymbol{P}_{27}\,\vdots\,\boldsymbol{P}_{28}\,\vdots\,\boldsymbol{P}_{29}]=[0\,\vdots\,2\;0\,\vdots\,0\,\vdots\,0\,\vdots\,0\;0\;0\,\vdots\,-2\;\;-2\,\vdots\,0]. \tag{67}
$$

此外，电阻、电容、电感的电压-电流特性式可以写成如下形式

$$
\begin{bmatrix} v_{R_1}(t) \\ v_{R_2}(t) \end{bmatrix}=\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} i_{R_1}(t) \\ i_{R_2}(t) \end{bmatrix}
$$

$$
\left[\begin{array}{c} i_{C_1}(t) \\ i_{C_2}(t) \\ i_{C_3}(t) \\ \hdashline i_{C_4}(t) \end{array}\right]=\left[\begin{array}{ccc:c} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ \hdashline 0 & 0 & 0 & 4 \end{array}\right]\left[\begin{array}{c} \dot{v}_{C_1}(t) \\ \dot{v}_{C_2}(t) \\ \dot{v}_{C_3}(t) \\ \hdashline \dot{v}_{C_4}(t) \end{array}\right]
$$

$$
\left[\begin{array}{c} v_{L_1}(t) \\ \hdashline v_{L_2}(t) \\ v_{L_3}(t) \end{array}\right]=\left[\begin{array}{c:cc} 1 & 0 & 0 \\ \hdashline 0 & 2 & 1 \\ 0 & 1 & 3 \end{array}\right]\left[\begin{array}{c} i_{L_1}(t) \\ \hdashline i_{L_2}(t) \\ i_{L_3}(t) \end{array}\right]
$$

将这些式子与(50)、(51)、(52)式比较,得

$$\left.\begin{aligned}
&\boldsymbol{R}_R=\begin{bmatrix}1 & 0\\ 0 & 2\end{bmatrix}\\
&\boldsymbol{C}_C=\begin{bmatrix}1 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 3\end{bmatrix}\qquad \boldsymbol{C}_S=4\\
&\boldsymbol{L}_{\Gamma\Gamma}=1\qquad\qquad \boldsymbol{L}_{LL}=\begin{bmatrix}2 & 1\\ 1 & 3\end{bmatrix}
\end{aligned}\right\}\tag{68}$$

将上面得到的(66)、(67)、(68)式代入(56)式,得

$$\left[\begin{array}{c|cc|c|c|ccc|cc}
1 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ \hline
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ \hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ \hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ \hline
0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0\\
0 & 0 & -1 & -4 & 0 & 0 & 2 & 0 & 0 & 0\\
1 & -1 & 1 & 4 & 0 & 0 & 0 & 3 & 0 & 0\\ \hline
0 & 0 & 0 & 0 & 1 & 0 & 0 & & 2 & 1\\
0 & 0 & 0 & 0 & 1 & 0 & 0 & & 1 & 3
\end{array}\right]
\left[\begin{array}{c}
v_{I_s}(t)\\ \hline
i_{R_1}(t)\\
i_{R_2}(t)\\ \hline
\dot{v}_{C_4}(t)\\ \hline
\dot{i}_{L_1}(t)\\ \hline
\dot{v}_{C_1}(t)\\
\dot{v}_{C_2}(t)\\
\dot{v}_{C_3}(t)\\ \hline
\dot{i}_{L_2}(t)\\
\dot{i}_{L_3}(t)
\end{array}\right]$$

$$=\left[\begin{array}{c|c|ccc|cc}
0 & 0 & 0 & 0 & 0 & -2 & -2\\ \hline
0 & 0 & -1 & 0 & -1 & 0 & 0\\
0 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} & 0 & 0\\ \hline
-1 & 0 & 0 & -1 & 1 & 0 & 0\\ \hline
0 & -1 & 0 & 0 & 0 & 1 & 1\\ \hline
0 & 0 & 0 & 0 & 0 & -1 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & -1 & -1\\ \hline
0 & 0 & 1 & 0 & 1 & 0 & 0\\
0 & 0 & 0 & 0 & 1 & 0 & 0
\end{array}\right]
\left[\begin{array}{c}
v_{C_4}(t)\\ \hline
i_{L_1}(t)\\ \hline
v_{C_1}(t)\\
v_{C_2}(t)\\
v_{C_3}(t)\\ \hline
i_{L_2}(t)\\
i_{L_3}(t)
\end{array}\right]
+\left[\begin{array}{c}
0\\ \hline
1\\
0\\ \hline
0\\ \hline
0\\ \hline
0\\
0\\
0\\ \hline
0\\
0
\end{array}\right]v_E(t)\tag{69}$$

因(69)式左边的系数矩阵的第 4 行和第 5 行是零行向量,所以不是正则矩阵. 因此,若将(69)式的第 4 行和第 5 行提出来改写成

$$-v_{C_4}(t)-v_{C_2}(t)+v_{C_3}(t)=0$$

$$-i_{L_1}(t)+i_{L_2}(t)+i_{L_3}(t)=0$$

将两式微分，得

$$\dot{v}_{C_4}(t) = -\dot{v}_{C_2}(t) + \dot{v}_{C_3}(t)$$

$$\dot{i}_{L_1}(t) = \dot{i}_{L_2}(t) + \dot{i}_{L_3}(t)$$

再将其代入(69)式中除了第 4 式、第 5 式以外的各式，则得

$$\left[\begin{array}{c|ccc|ccc|cc}
1 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & -1 & 0 & 6 & -4 & 0 & 0 \\
1 & -1 & 1 & 0 & -4 & 7 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 3 & 2 \\
0 & 0 & 0 & 0 & 0 & 0 & 2 & 4
\end{array}\right]
\left[\begin{array}{c}
i_{I_d}(t) \\ \hline
i_{R_1}(t) \\ i_{R_2}(t) \\ \hline
\dot{v}_{C_1}(t) \\ \dot{v}_{C_2}(t) \\ \dot{v}_{C_3}(t) \\ \hline
\dot{i}_{L_2}(t) \\ \dot{i}_{L_3}(t)
\end{array}\right]$$

$$= \left[\begin{array}{ccc|cc}
0 & 0 & 0 & -2 & -2 \\
\hline
-1 & 0 & -1 & 0 & 0 \\
0 & -\frac{1}{2} & \frac{1}{2} & 0 & 0 \\
\hline
0 & 0 & 0 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & -1 \\
\hline
1 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0
\end{array}\right]
\left[\begin{array}{c}
v_{C_1}(t) \\ v_{C_2}(t) \\ v_{C_3}(t) \\ \hline
i_{L_2}(t) \\ i_{L_3}(t)
\end{array}\right]
+ \left[\begin{array}{c}
0 \\ \hline 1 \\ 0 \\ \hline 0 \\ 0 \\ 0 \\ \hline 0 \\ 0
\end{array}\right] v_E(t) \qquad (70)$$

因(70)式左边系数矩阵是正则的，则由基本行变换可变换成单位矩阵．最后结果是

$$\begin{bmatrix}
i_{I_d}(t) \\ i_{R_1}(t) \\ i_{R_2}(t) \\ \dot{v}_{C_1}(t) \\ \dot{v}_{C_2}(t) \\ \dot{v}_{C_3}(t) \\ \dot{i}_{L_2}(t) \\ \dot{i}_{L_3}(t)
\end{bmatrix}
= \begin{bmatrix}
-2 & 0 & -2 & -2 & -2 \\
-1 & 0 & -1 & 0 & 0 \\
0 & -\frac{1}{2} & \frac{1}{2} & 0 & 0 \\
-1 & 0 & -1 & -1 & 0 \\
\frac{2}{13} & -\frac{3}{52} & \frac{11}{52} & \frac{2}{13} & \frac{2}{13} \\
\frac{3}{13} & \frac{1}{26} & \frac{5}{26} & \frac{3}{13} & \frac{3}{13} \\
\frac{1}{2} & 0 & \frac{1}{4} & 0 & 0 \\
-\frac{1}{4} & 0 & \frac{1}{8} & 0 & 0
\end{bmatrix}
\begin{bmatrix}
v_{C_1}(t) \\ v_{C_2}(t) \\ v_{C_3}(t) \\ i_{L_2}(t) \\ i_{L_3}(t)
\end{bmatrix}
+ \begin{bmatrix}
2 \\ 1 \\ 0 \\ 1 \\ -\frac{2}{13} \\ -\frac{3}{13} \\ 0 \\ 0
\end{bmatrix} v_E(t) \qquad (71)$$

由(71)式的下半部分可以得出下列形式的标准形微分方程式

$$
\begin{bmatrix} \dot{v}_{C_1}(t) \\ \dot{v}_{C_2}(t) \\ \dot{v}_{C_3}(t) \\ \dot{i}_{L_1}(t) \\ \dot{i}_{L_2}(t) \end{bmatrix}
=
\begin{bmatrix}
-1 & 0 & -1 & -1 & 0 \\
\frac{2}{13} & -\frac{3}{52} & \frac{11}{52} & \frac{2}{13} & \frac{2}{13} \\
\frac{3}{13} & \frac{1}{26} & \frac{5}{26} & \frac{3}{13} & \frac{3}{13} \\
\frac{1}{2} & 0 & \frac{1}{4} & 0 & 0 \\
-\frac{1}{4} & 0 & \frac{1}{8} & 0 & 0
\end{bmatrix}
\begin{bmatrix} v_{C_1}(t) \\ v_{C_2}(t) \\ v_{C_3}(t) \\ i_{L_1}(t) \\ i_{L_2}(t) \end{bmatrix}
+
\begin{bmatrix} 1 \\ -\frac{2}{13} \\ -\frac{3}{13} \\ 0 \\ 0 \end{bmatrix} v_E(t) \qquad (72)
$$

输出也可以直接由(71)式求出

$$
\text{输出} = i_{I_6}(t) = [-2\ 0\ -2\ -2\ -2] \begin{bmatrix} v_{C_1}(t) \\ v_{C_2}(t) \\ v_{C_3}(t) \\ i_{L_1}(t) \\ i_{L_2}(t) \end{bmatrix} + 2v_E(t)
$$

(问题 3) 试证明在满足假定 A(i)～A(vi) 的网络中，应用上述关于一般非时变线性网络状态方程式的推导方法，可以得到(31)式.

(问题 4) 在满足假定 A(i), (ii), (vii) 的 RLCM 网络中，试将用上述关于一般非时变线性网络状态方程式的推导方法得到的对应于(64)式的标准形微分方程式与(41)式进行比较，将对应于(64)式中 $\boldsymbol{x}_{2n}$ 的变量与(40)式中的 $\boldsymbol{q}$、$\boldsymbol{\phi}$ 及(40)′式中的 $\boldsymbol{e}$、$\boldsymbol{j}$ 进行比较. 其中假定，由补充状态变量 $\boldsymbol{x}_0$ 中取掉用其它分量和输入的线性组合表示的分量得到准状态变量 $\boldsymbol{x}_n$ 时，是取掉 $\boldsymbol{V}_S$ 和 $\boldsymbol{I}_\Gamma$.

A-I-3 线性物理系统的状态方程式

引言

在 **A-I-2** 中讲过的非时变线性网络状态方程式的推导方法，大致可分为下列五步：

S1 规定各元件电压、电流的基准方向，画出具有与该基准方向相对应枝的方向的有向图.

S2 找出该图的 N 树，写出基本割集方程式和基本回路方程式.

S3 利用电阻及从属电源的电压-电流特性式，消去基本割集方程式、基本回路方程式中的电阻和从属电源上的电压、电流.

S4 在这样得到的方程式中，再代入电容及电感的电压-电流特性式，消去电容电流和电感电压.

S5 在电容电压、电感电流中，消去非独立的分量，即消去可以用其它电容电压、电感电流或独立电压源电压、独立电流源电流的线性组合表示的分量.

本节的目的是要说明，该方法也适用于除了电气网络以外的其它物理系统. 为此，必须考虑下列三个问题.

Q1 是否存在象电气网络中的电压、电流一样，满足克希霍夫第一、第二定律的一一对应的物理量.

Q2 系统中组成环节的特性能否依据这些物理量用线性代数方程式（象电阻一样）或一阶线性微分方程式（象电容、电感一样）表示.

Q3 是否可以规定这些物理量的基准方向，将系统的组成画成有向图.

若对于某物理系统上述三点均可能，则该物理系统和电气网络在数学上具有完全相同的性质. 利用和电气网络中完全相同的方法，可以推导出该物理系统的状态方程式. 下面将要讨论的 1 维机械系统（直线运动系统、回转运动系统）、流体系统及电气-机械系统等，就是这种物理系统的例子. 在这些物理系统中，满足克希霍夫第一定律的物理量均叫做通过变量（through variable），满足克希霍夫第二定律的物理量均称为横断变量（across variable）. 在 A-I-2 电气网络状态方程式的推导方法中，若将电压换成横断变量，电流换为通过变量，则该方法除了适用于电气网络以外，也适用于其它物理系统. 因此，以下将重点说明，在所讨论的物理系统中，哪些物理量是横断变量，哪些物理量是通过变量. 为了避免重复，状态方程式的推导方法不再改述. 此外，为了说明简单，假定所讨论系统中各环节的特性是线性的，而且不随时间变化.

横断变量和通过变量[45~47]

我们先来说明，在一维机械系统（直线运动系统、回转运动系统）、流体系统中哪些物理量是横断变量，哪些物理量是通过变量.

首先讨论直线运动系统中的横断变量. 在直线运动系统中，与该系统运动方向平行地设置固定在大地上的直线坐标轴. 这时，任何端子在该坐标系统中的坐标称为该端子

的位移，两个端子（例如端子 ⓐ 和端子 ⓑ）的位移 x_a、x_b 之差定义为两个端子间的位移差 δ，即

$$\delta = x_b - x_a \tag{1a}$$

定义为以 ⓐ 的位移为基准的 ⓐ、ⓑ 间的位移差；

$$\delta = x_a - x_b \tag{1b}$$

定义为以 ⓑ 的位移为基准的 ⓐ、ⓑ 间的位移差．由(1a)、(1b)两式可见，以 ⓐ 的位移为基准时 ⓐ、ⓑ 之间的位移差和以 ⓑ 的位移为基准时 ⓐ、ⓑ 之间的位移差，绝对值相等，而符号相反．这个事实说明，可以规定位移差的基准方向．因此规定，以 ⓐ 的位移为基准讨论 ⓐ、ⓑ 之间的位移差时，图中枝的方向标成由 ⓑ 到 ⓐ(图 1(a))；以 ⓑ 的位移为基准讨论 ⓐ、ⓑ 之间的位移差时，图中枝的方向标成由 ⓐ 到 ⓑ(图 1(b))．以上定义位移差的方法对应于，在电气系统中可以根据以大地为基准的各端子的电位之差定义两端子间的电压．由电压满足克希霍夫第二定律类推，显然上面定义的位移差也满足克希霍夫第二定律．因此，以上定义的位移差是直线运动系统的横断变量．

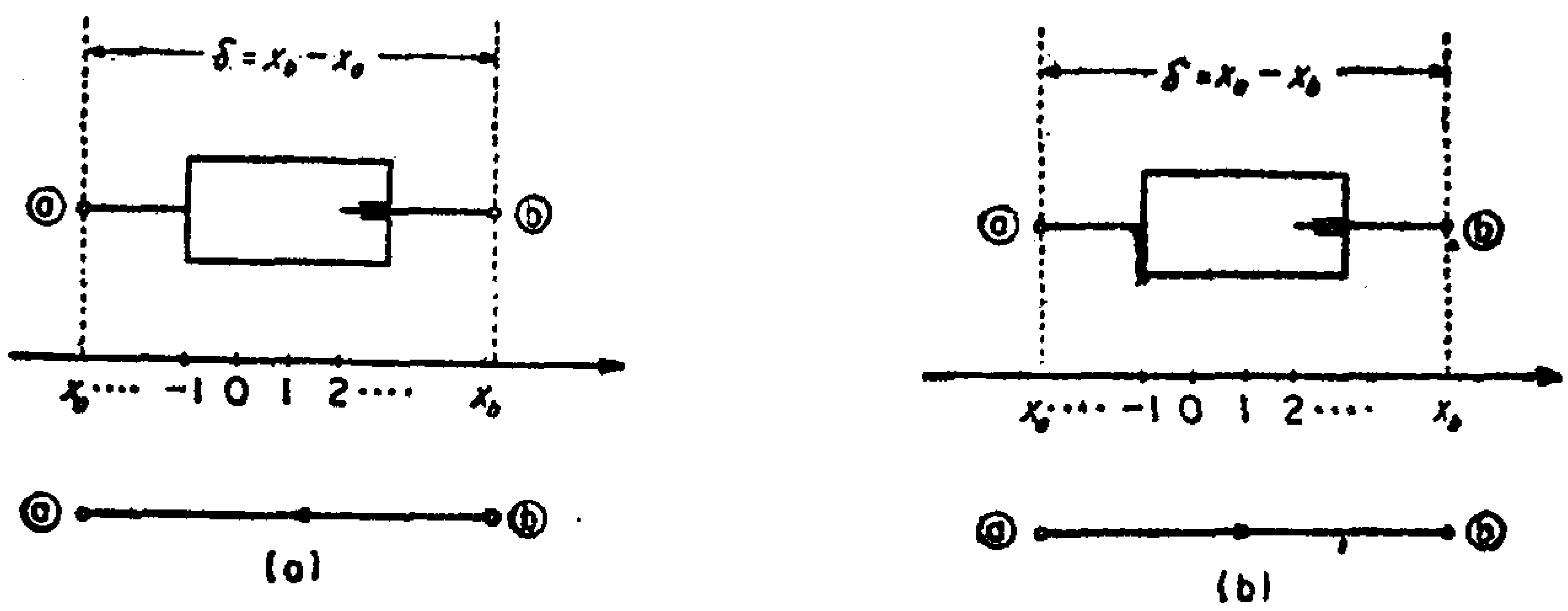
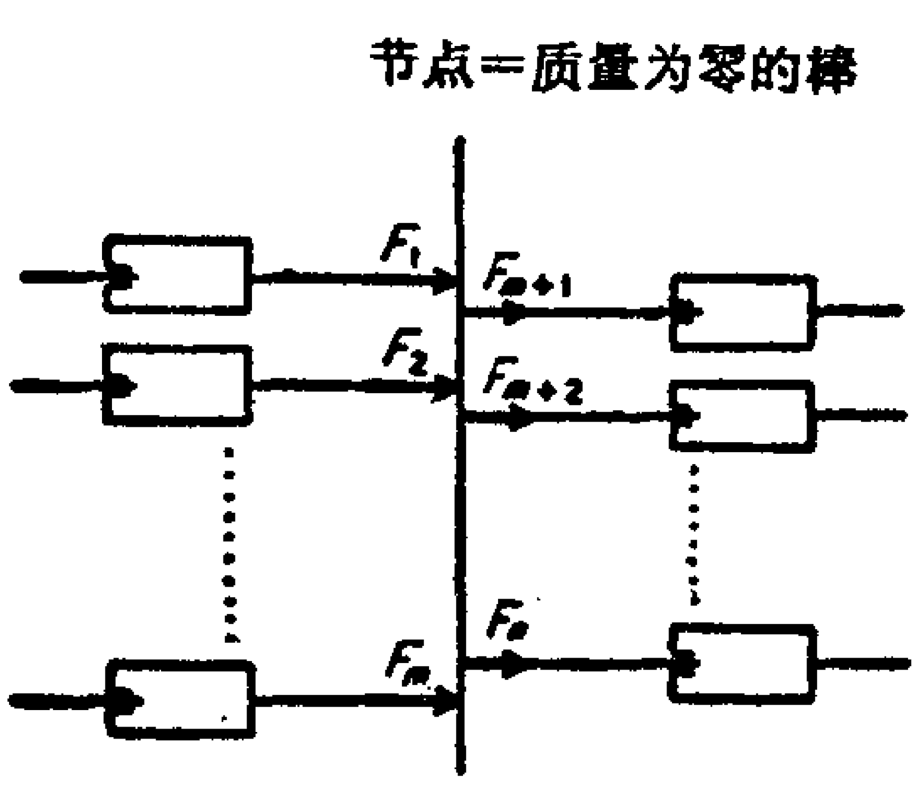


图 1　位移差的基准方向和图中枝的方向之间的关系

其次，再来谈一下直线运动系统的通过变量．在直线运动系统中，假定上面设置的静止坐标轴的正方向为力 F 的基准方向．根据众所周知的力学中的作用和反作用定律，这时作用在任意节点[1]（环节的结合点）上力的总和是零（参看图 2)．亦即，作用在节点上的力满足克希霍夫第一定律．因此，具有按上述方法规定基准方向的力，是直线运动系统的通过变量．



$$(F_1+F_2+\cdots+F_m)+(F_{m+1}+F_{m+2}+\cdots+F_n)=0$$

图 2　直线运动系统中节点上力的关系

（注意）　克希霍夫第一定律本来是物质在流动过程中必须遵守的物理定律．因为电气系统中的通过变量电流实质上是一种流动，所以电流满足克希霍夫第一定律是完全可以理解的．与此相应，力不是物质流动，为了将它看成通过变量，如上所述，必须考虑对于所有力共同的基准方向．

上面讨论了环节的位移差和图中枝的方向的关系，现在我们再来看一下作用于环节上的力与有向图中枝的方向有什么关系．因为力不是物质流动，所以在对直线运动系统描绘出的图中，枝中所标方向的意义，与电气系统中电流方向的意义不

1) 如图 2 所示，实际上直线运动系统的节点多半是线．

同．亦即，在电气系统中与电流这个通过变量的关系上，有向图中枝的方向表示电流流动的基准方向，而在直线运动系统中，象对于缓冲器及弹簧等 2 个端子的环节，枝的方向表示外力是作用在哪一端．如下面例子所示，在直线运动系统中必须指出外力是作用在系统的哪端．其理由是，若仅指出作用于环节上的外力（大小及方向），而不说明该外力是作用在哪一端，便不能唯一地确定环节的特性．

试看一下质量为零的理想化的弹簧．设其 2 个端子为ⓐ、ⓑ．假定我们只是说“外力 F 作用在弹簧上”．如图 3(a)、(b)所示，我们来看一下 F 作用在 ⓐ（图 3(a)）和作用在 ⓑ 的两种情况．上述的“外力 F……”并没有指明是其中的哪一种情况[47]．

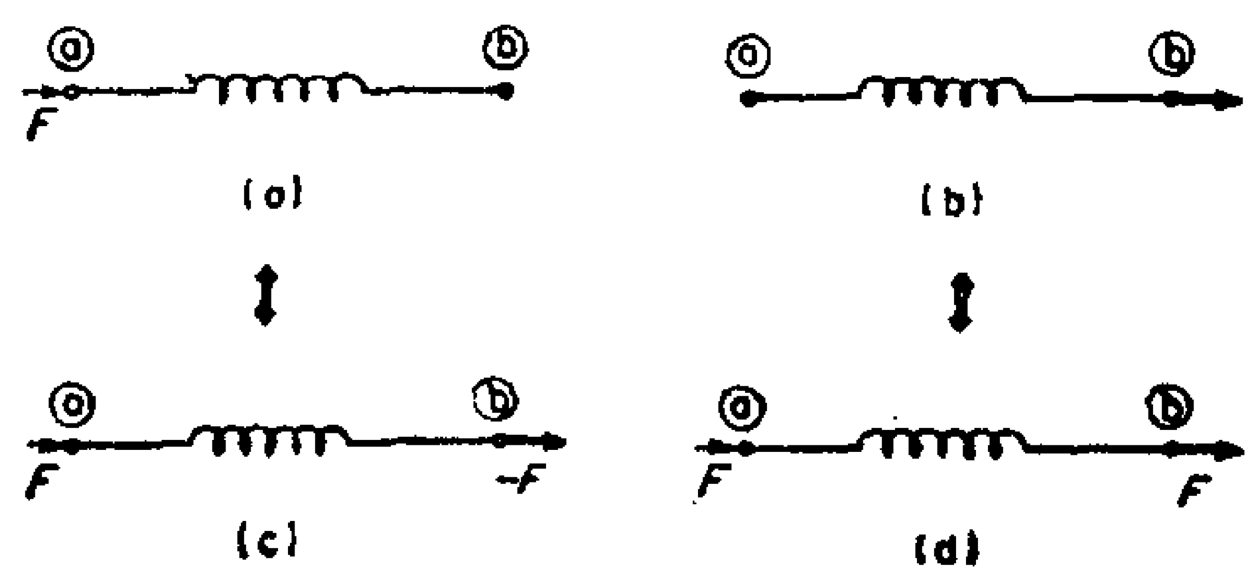


图 3　作用在弹簧上的力

在弹簧这样的质量为零的环节上，根据作用与反作用定律，两个端子上外力的绝对值相等，而符号相反．亦即，外力 F 作用在端子 ⓐ 时，在端子 ⓑ 上作用着外力 $-F$（图 3(c)）．同样，图 3(b) 和图 3(c) 表示着完全相同的状态．由图 3 可见，当正的力作用在端子 ⓐ 时，使弹簧缩短；而当正的力作用在端子 ⓑ 时，使弹簧伸长．因此，即使知道作用于环节上的外力（大小及方向），但若不指出该外力是作用在那一端，仍然不能唯一地确定该环节的特性．

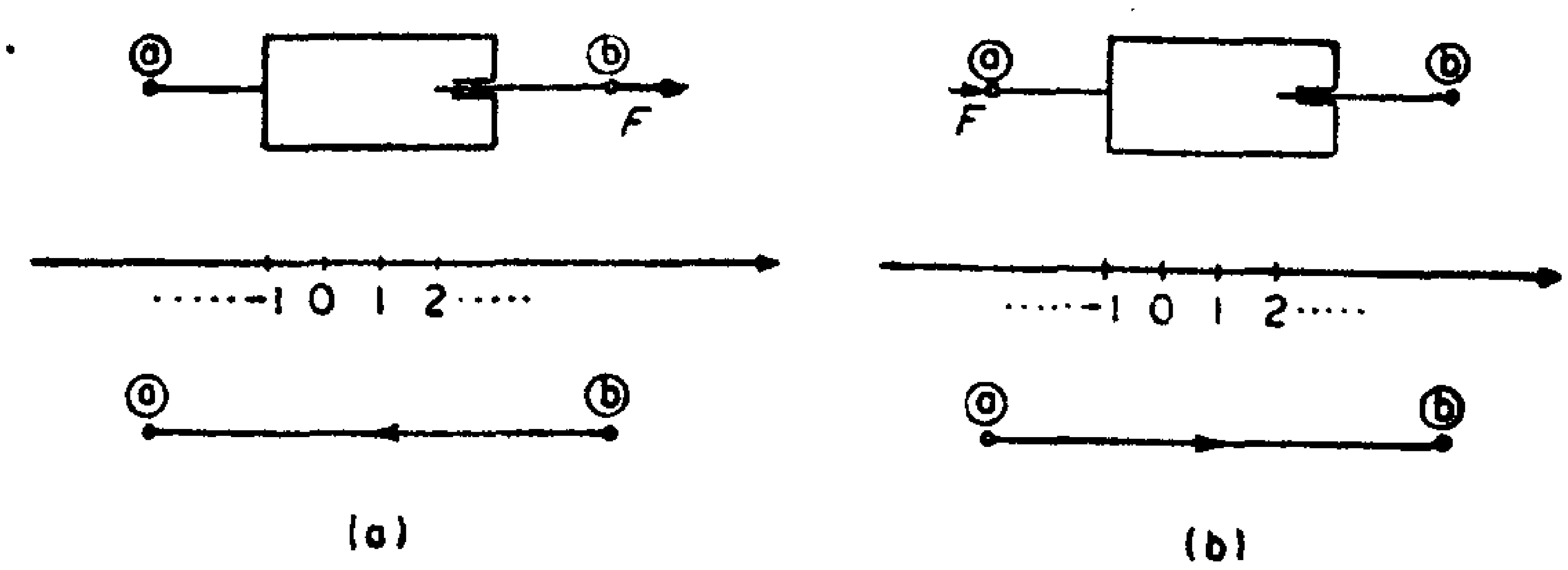


图 4　标明外力的方法及其与图中枝的关系

因此规定，用枝的方向表示外力是作用在哪一端．亦即，在具有端子 ⓐ、ⓑ 的 2 端子环节上，当指明外力是作用在端子 ⓑ 时，如图 4(a)所示，枝的方向标成由 ⓑ 到 ⓐ；当指明外力是作用在端子 ⓐ 时，如图 4(b)所示，枝的方向标成由 ⓐ 到 ⓑ．根据该结论，在弹簧及缓冲器这样的质量视为零的理想化的 2 端环节上，根据作用和反作用定律，因两个端子上外力的大小相等，方向相反，所以改变对应于该环节图中枝的方向，就相当于改变规定了基准方向的外力的符号．在电气系统的图中也可以作如下解释．由图 5(a) 可见，枝中标以由 ⓑ 到 ⓐ 的方向表示指出电流的方向是由端子 ⓑ 流入环节．同样，枝中标以由 ⓐ 到 ⓑ 的方向表示指出电流的方向是由端子 ⓐ 流入环节（参看图 5(b)）．若这样解释电气系统图中枝的方向和电流（电气系统中的通过变量）的关系，则上述规定的直线运动系统图中枝的方向和力（直线运动系统中的通过变量）的关系也可以由电气系统类推．

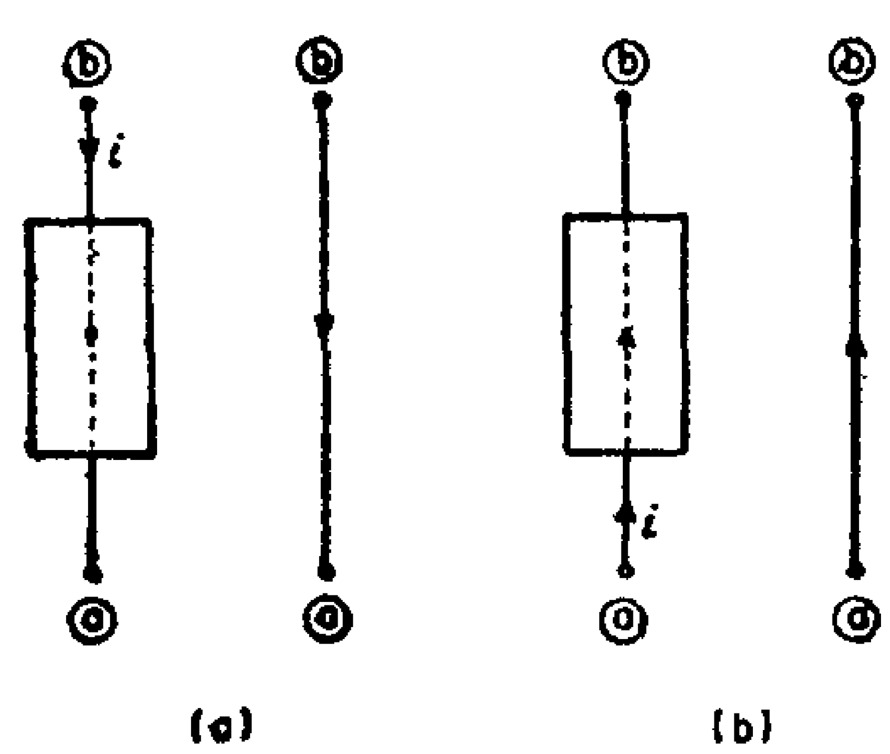


图 5　电气系统中流入环节的电流和图中枝的方向

以上，分别讨论了位移差的基准方向和枝的方向的对应关系，并且指出了外力作用在环节的哪一端用枝的方向表示的问题．但是，对于一个系统，位移差的有向图和力的有向图各自存在，将一个系统用一个有向图表示是不方便的．因此，象是在电气系统中端电压与端电流相互关联一样，位移差的基准方向和指出外力作用在哪一端有如下关系．亦即，在具有端子 ⓐ、ⓑ 的 2 端环节上，以 ⓐ 的位移为基准考察 ⓐ、ⓑ 之间的位移差时，规定指明的是 ⓑ 上的外力．这时由图 1(a) 和图 4(a) 可见，图中枝的方向标成由 ⓑ 指向 ⓐ．反之，以 ⓑ 的位移为基准考察 ⓐ、ⓑ 之间的位移差时，规定指明的是 ⓐ 上的外力．图中枝的方向标成由 ⓐ 指向 ⓑ．注意，象这样地将位移差和外力联系起来，则位移差的微分和外力的乘积就是单位时间内供给环节的能量．

其次，再来讨论一下回转运动系统的横断变量和通过变量．在该系统中，设系统的回转运动轴为圆柱轴，坐标系统是固定在大地上的圆柱坐标系统．任意端子在该坐标系统中的角坐标为该端子的角位移 ψ，两个端子之间角位移差为端子间的角位移差 φ．亦即，当端子 ⓐ、ⓑ 的角位移分别为 ψ_a、ψ_b 时，

$$\varphi=\psi_b-\psi_a \tag{2a}$$

定义为以 ⓐ 的角位移为基准时 ⓐ、ⓑ 间的角位移差，而

$$\varphi=\psi_a-\psi_b \tag{2b}$$

定义为以 ⓑ 的角位移为基准时 ⓐ、ⓑ 间的角位移差．将这种角位移差的定义与直线运动系统中位移差的定义相比较，显然两者完全是一一对应的．因此，与“位移差是直线运动系统的横断变量”相对应，可以说回转角位移差是回转系统的横断变量．

另外，若将上述设置的静止圆柱坐标系统角坐标的正方向取为力矩 T 的基准方向，则象力是直线运动系统的通过变量一样，力矩是回转运动系统的通过变量．在直线运动系统中，对于作用在象缓冲器及弹簧这样 2 端环节上的外力，不仅要指出其大小和方向，而且必须指出该外力是作用在环节的哪一端．与此类似，在回转运动系统中，对于作用在缓冲器及弹簧上的力矩，还必须指出该力矩是作用在哪一端．

在直线运动系统的图中，枝的方向表示位移差的基准方向和外力是作用在哪一端．与此类似，回转系统也可以用图表示，在该图中枝的方向表示出回转角位移差的基准方向和力矩是作用在哪一端．象在电气系统及直线运动系统中一样，在回转系统中，为了将一个系统用一个图表示，也将回转角位移差的基准方向和力矩是作用在哪一端联系起来，枝的方向作如下规定．亦即，在具有端子 ⓐ、ⓑ 的 2 端环节上，当以 ⓐ 的回转角位移为基准讨论 ⓐ、ⓑ 之间的回转角位移差时，规定是指出作用在 ⓑ 端的力矩，这时图中枝的方向是由 ⓑ 指向 ⓐ(参看图 6(a))．反之，当以 ⓑ 的回转角位移为基准讨论 ⓐ、ⓑ 之间的回转角位移差时，规定是指出作用在 ⓐ 端的力矩，这时图中枝的方向是由 ⓐ 指向 ⓑ(参看图 6(b))．若将回转角位移差和力矩这样联系起来，则回转角位位移差的微分和力矩的乘积便是单位时间内供给环节的能量．

最后，我们再来讨论一下流体系统的横断变量和通过变量．显然，在流体系统中，若将压力差及流量分别和电气系统中的电压(电位差)及电流进行对比，则两个端子间的压力差 P 满足克希霍夫第二定律，流量 f 满足克希霍夫第一定律．亦即，压力差是流体系统的横断变量，流量是该系统的通过变量．在流体系统中也象在电气系统、直线运动系统和回转系统中一样，为了将一个系统用一个有向图表示，必须将压力差的基准方向和流量

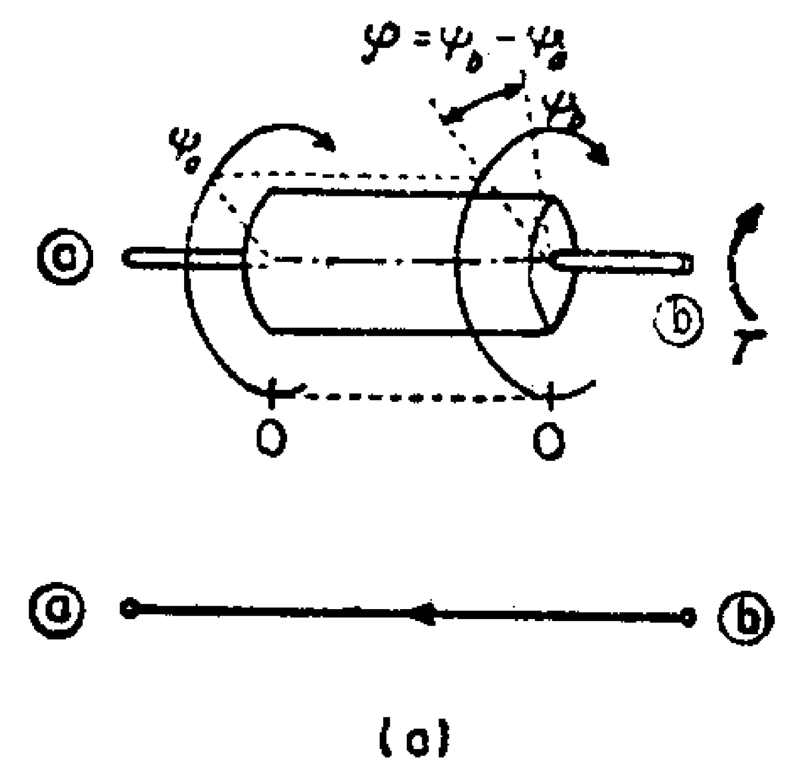


图 6　回转角位移差 φ、力矩 T 及枝的方向之间的关系

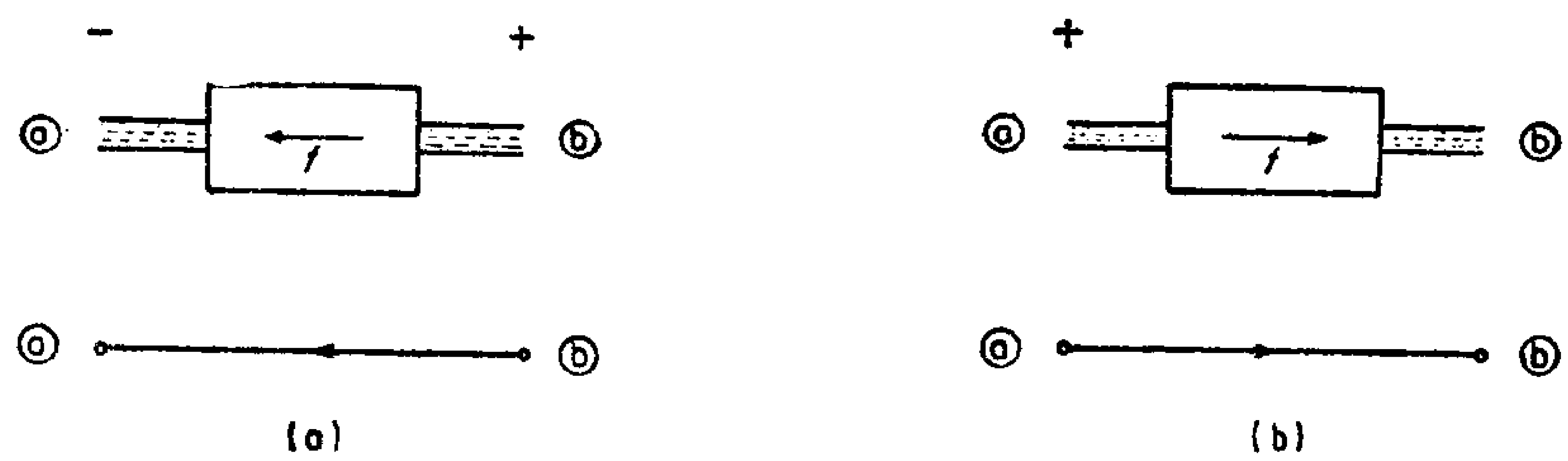


图 7　压力差、流量的基准方向与图中枝的方向

的基准方向联系起来，规定按下列规则标出枝的方向．亦即，在具有端子 ⓐ、ⓑ 的 2 端环节上，当以 ⓐ 的压力为基准讨论 ⓐ、ⓑ 之间的压力差时，规定流量的基准方向取为由 ⓑ 到 ⓐ，而枝的方向标成由 ⓑ 到 ⓐ(图 7(a))．反之，当以 ⓑ 的压力为基准讨论 ⓐ、ⓑ 之间的压力差时，则规定流量的基准方向取为由 ⓐ 到 ⓑ．而枝的方向标成由 ⓐ 到 ⓑ(图 7(b))．

以上，我们分别在直线运动系统、回转运动系统及流体运动系统中了解到具有横断变量及通过变量性质的物理量．因为克希霍夫第一定律和第二定律在任意时刻均成立，若某个变量(物理量)满足克希霍夫第一定律或第二定律，则该变量的微分及积分也满足克希霍夫第一定律或第二定律．例如，横断变量的微分及积分也是横断变量．因此，以下将位移差 δ 的微分，即速度差 $v(\triangleq\delta')$ 作为直线运动系统的横断变量；回转角位移差 φ 的微分，即角速度差 $\omega(\triangleq\dot{\varphi})$ 作为回转系统的横断变量．为了将直线运动系统及回转运动系统基本环节的端子特性用(线性)代数方程式或一阶(线性)微分方程式表示(这样表示基本环节的端子特性，便于采用和电气系统统一的处理方法)，所以必须用 v 和 ω 代替 δ 及 φ．将各系统的横断变量和通过变量整理后，示于表 1．

表 1　各系统的通过变量及横断变量

系　统	通过变量	x	横断变量	y
电气系统	电流	i	电压	v
直线运动系统	力	F	速度差	v
回转运动系统	力矩	T	角速度差	ω
流体运动系统	流量	f	压力差	P

直线运动系统、回转运动系统、流体系统的基本环节[45-48]

为了使电气系统中状态方程式的推导方法同样适用于本节所讨论的直线运动系统，必须绘出这些系统的图，并且象电气系统中的电阻、电容、电感等的特性一样，必须将组成系统的各环节的特性用横断变量及通过变量的端子特性表示. 为了绘出有向图，并且将环节的特性用端子特性式表示，首先必须找出环节的端子对. 在电气系统的环节上，象电阻、电容、电感等，可以直接由其外形决定端子对. 但是在其它物理系统中，往往不能这样直观地决定环节的端子对.

在直线运动系统中，象图 8(a) 弹簧及图 8(b) 缓冲器的端子对，可以很容易地由其形状找出. 弹簧的端子特性是

$$v(t)=\frac{1}{k}\dot{F}(t) \tag{3}$$

而缓冲器的端子特性可以用

$$F(t)=bv(t) \tag{4}$$

表示. 若如上述那样地将位移差(因而速度差)的基准方向和指出外力是作用在哪一端联系起来，则无论枝的方向如何，在视为无源环节的一般弹簧及缓冲器中，k, b 都是正值.

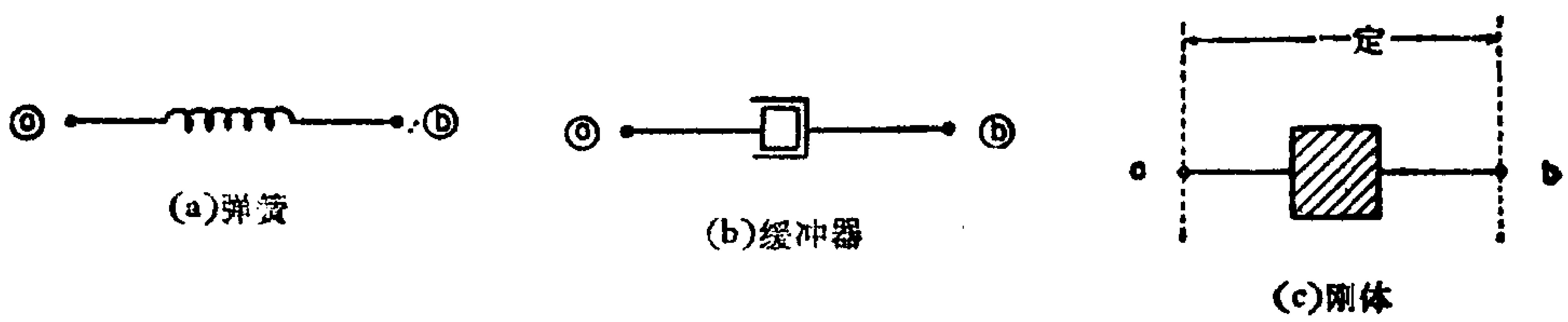


图 8　直线运动系统的基本环节

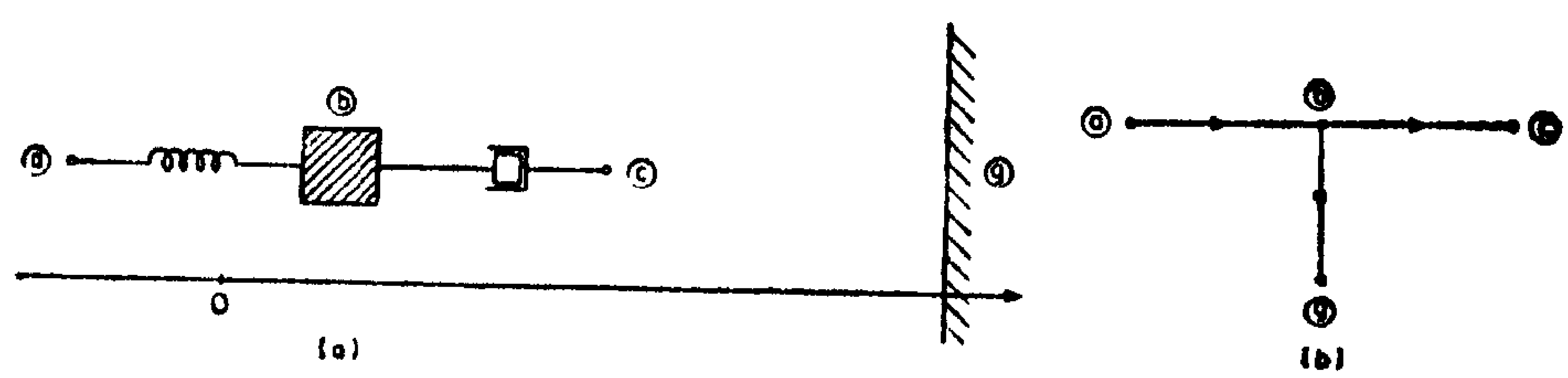


图 9　包含有刚体的直线运动系统的例子及其图

与此相应，在直线运动系统中还有一个重要的基本环节——刚体(图 8(c)). 若根据刚体的形状将 a、b 取为端子对，因 a、b 之间的速度差始终为零，所以该速度差不能表示环节的特性. 因此，如下所述，必须根据刚体的物理性质确定其端子对. 亦即，因为在直线运动中刚体可以作为具有与其相同质量的质点一样地处理，其特性可以用外力和绝对加速度[1)]来表示，所以将一个端子取在刚体本身的适当点(例如重心)，而另一个取在前述静止坐标轴上适当的固定点(例如坐标原点)，即在表示刚体的图枝上，其中的一个节点总是固定点. 因此，与刚体相连接的环节都必须看成是与刚体自身上那个端子相连. 例如，图 9(a) 所示系统可以用 (b) 的图表示. 这时从物理意义看，因为只能指出作用在刚体自

1) 绝对加速度表示坐标系统固定在大地时的加速度.

身侧的那个端子上的外力，所以枝的方向必须标成由该节点指向固定点的节点[1]. 象这样确定端子对时，刚体的端子特性可以表示成

$$F(t)=m\dot{v}(t) \tag{5}$$

若将位移差(因而速度差)的基本方向和外力是作用在哪一端按前述规定联系起来，则在视为无源环节的一般刚体上，m 是正值.

对于回转运动系统中的弹簧、缓冲器及刚体，可以说也和直线运动系统的情况一样. 首先，图 10(a)的弹簧和(b)的缓冲器的端子对可由其形状找出，弹簧的端子特性为

$$\omega(t)=\frac{1}{k}\dot{T}(t) \tag{6}$$

缓冲器的端子特性可以表示成

$$T(t)=b\omega(t) \tag{7}$$

和直线运动系统的情况一样，若将回转角位移差(因而角速度差)的基准方向和力矩是作用在哪一端按前述规定联系起来，则无论图中枝的方向如何，在看作无源环节中的一般弹簧及缓冲器上，k、b 均为正值.

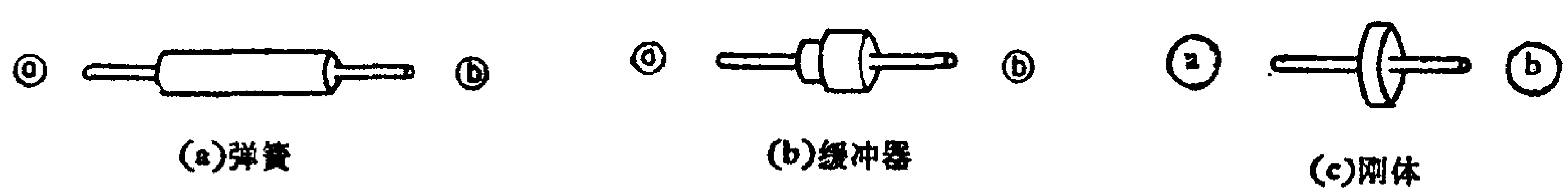


图 10 回转运动系统中的基本环节

在回转运动系统中，表示刚体(图 10(c))特性的端子对，不能由其形状决定. 因此，在这种情况下，依然必须根据刚体的物理性质确定其端子对. 即在回转运动中，可以将刚体看成具有与其相同惯性矩，厚度为零的圆盘. 因其特性可以由力矩和绝对角加速度表示，所以把其中的一个端子选在刚体内的某个适当点，而另一个端子取为适当的固定点. 即在回转运动系统里，表示刚体的图枝中的一个节点始终是固定点. 因此，与刚体相连的环节都必须看成是和刚体自身侧的端子相连. 例如，图 11(a)所示系统可以用(b)的图表示. 这时，从物理意义上看，因为力矩只能作用在刚体自身侧的那个端子，所以枝的方向标成由该节点指向固定点的节点[2]. 当这样确定端子对时，刚体的端子特性可以用

$$T(t)=j\dot{\omega}(t) \tag{8}$$

表示. 若象前述那样，将回转角位移差(因而角速度差)的基准方向和力矩是作用在哪一端联系起来，则在视为无源环节的一般刚体中，j 的值是正的.

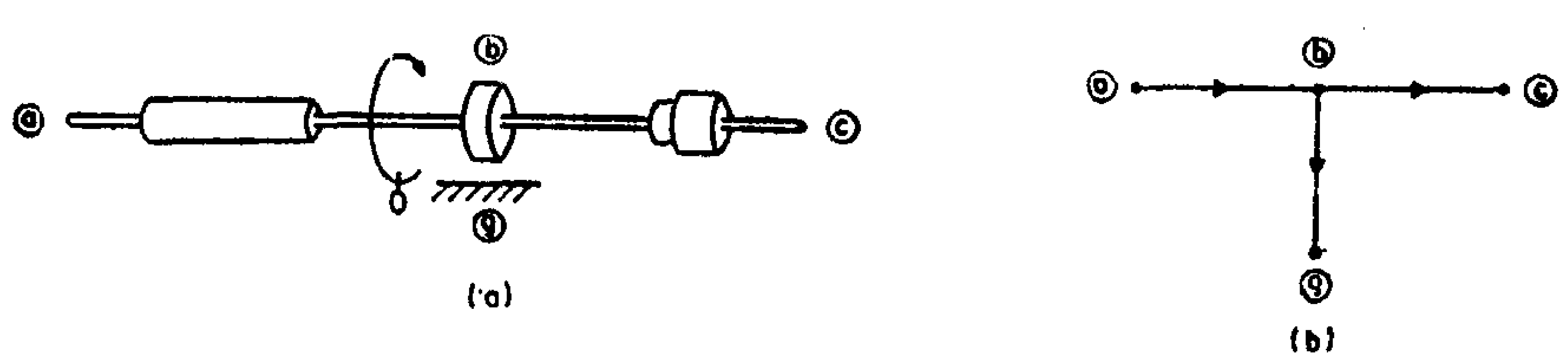
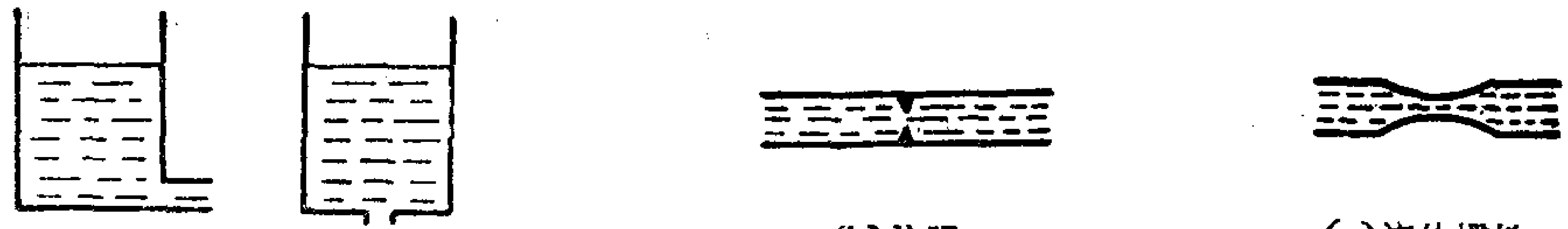


图 11 包含有刚体的回转运动系统的例子及其图

1) 从数学角度看，枝的方向标向哪一端都可以，若标成由固定点侧端子指向刚体侧端子，则作用在刚体上的力是 $-F$.

2) 和在直线运动系统中的情况一样，从数学角度看，无论枝的方向指向哪一端均可.



(a)贮水箱　(b)流阻　(c)流体惯性

图 12　流体系统的基本环节

现在，我们来讨论流体系统的基本环节，图 12(a)的贮水箱，(b)的流阻及(c)的流体惯性[48]．象直线运动系统及回转运动系统中的刚体一样，贮水箱的端子对也不能由其形状，而必须由其物理性质确定．亦即，如图 13 所示，若设贮水箱的流出口压力为 P，流量为 f，大气压力为 P_0，液面高度为 h，水箱截面积为 A，液体比重为 ρ，则下列关系成立

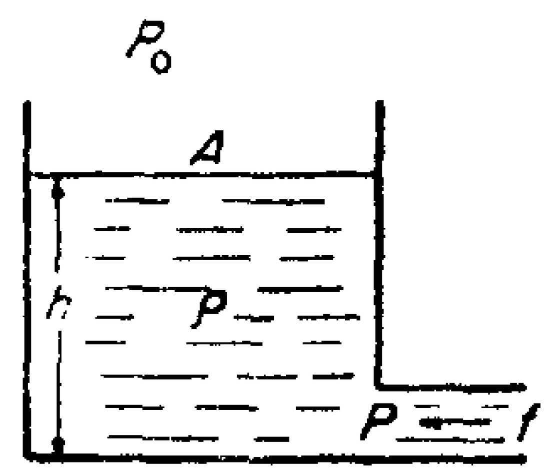


图 13　贮水箱

$$P(t)=P_0+\rho h(t)$$

$$\dot{h}(t)=\frac{1}{A}f(t).$$

由上列两式消去 h，得

$$\frac{d}{dt}(P(t)-P_0)=\frac{\rho}{A}f(t).$$

由该式可见，为了用流量及压力差表示贮水箱的端子特性，将其中的一个端子取在贮水箱的流出口，而另一个端子取在大气．这样确定端子对时，贮水箱的端子特性可以表示成

$$f(t)=c\dot{P}(t)\quad\left(c=\frac{A}{\rho}\right)\tag{9}$$

此外，流阻及流体惯性的端子对可由其物理形状确定，其端子特性可以表示成

$$P(t)=rf(t)\tag{10}$$

$$P(t)=l\dot{f}(t)\tag{11}$$

若象前述那样地将压力差和流量的基准方向联系起来，则与图中枝的方向如何标无关，在视为无源环节的一般贮水箱、流阻及流体惯性中，常数 c、r 及 l 的值是正的．

上述线性电气系统、机械系统及流体系统等当中的基本环节的共同特点是，其端子特性可以用线性代数方程式或一阶线性微分方程式表示．因此，为了便于统一地讨论这些物理系统，将这些基本环节按其共同特性进行分类．现设通过变量为 x，横断变量为 y，进行如下分类．

[定义 1][45]1)

(i) 端子特性用代数方程式

$$x(t)=\alpha_1 y(t)\quad \alpha_1\text{: 实常数}$$

表示的环节叫做代数环节(algebraic element)．

(ii) 端子特性用一阶线性微分方程式

$$x(t)=\alpha_2\dot{y}(t)\quad \alpha_2\text{: 实常数}$$

表示的环节叫做一阶横断环节(first-order across element)．

(iii) 端子特性用一阶线性微分方程式

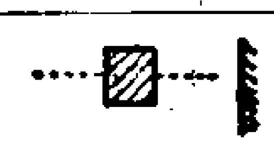
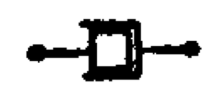
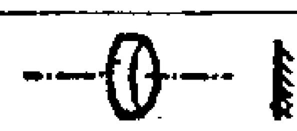
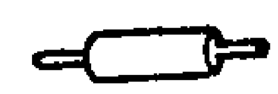
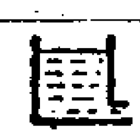
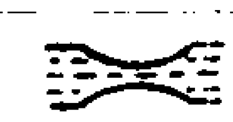
1) 在文献[49]中，分别将代数环节、一阶横断环节及一阶通过环节叫做广义电阻环节、广义电容环节及广义诱导环节．

$$y(t)=\alpha_3\dot{x}(t) \quad \alpha_3\text{: 实常数}$$

表示的环节叫做一阶通过环节(first-order through element).

基于表1确定的横断变量和通过变量，按上述定义将各系统的基本环节分类列于表2[1].

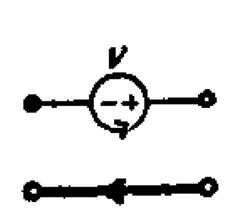
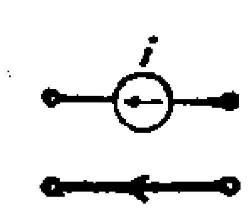
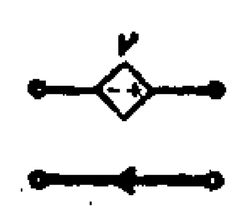
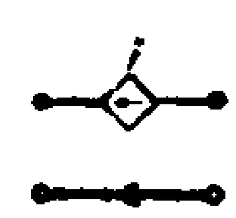
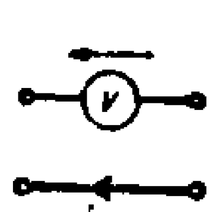
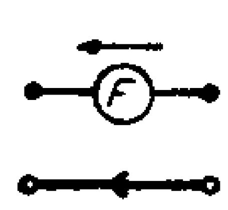
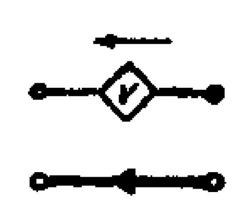
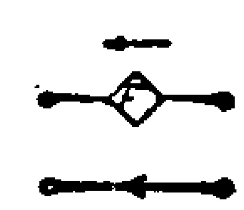
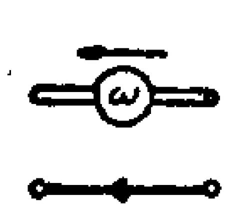
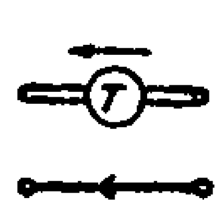
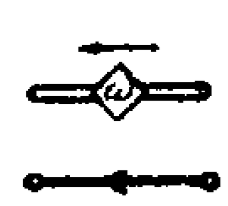
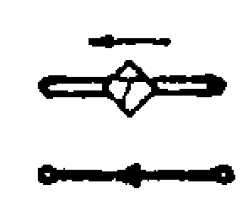
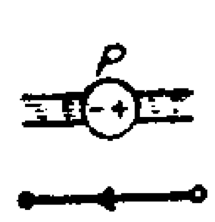
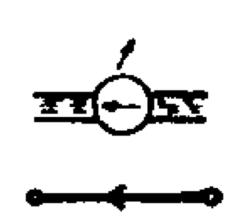
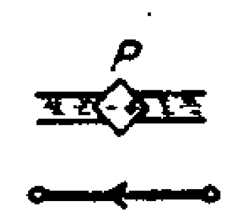
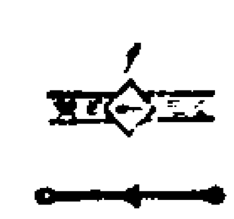
表 2

系统	一阶横断环节	代数环节	一阶通过环节
电气系统	 电容	电阻	 电感
直线运动系统	 刚体	 缓冲器	 弹簧
回转运动系统	 刚体	 缓冲器	 弹簧
流体系统	 贮水箱	 流阻	 流体惯性

象电气系统中的独立电压源、电流源及从属电压源、电流源那样，下面再来谈一下各系统的驱动源(source)及有源环节，以结束对基本环节的说明.

在直线运动系统中，存在着可以独立指定速度差(该系统的横断变量)的独立速度源，以及可以独立指定力(系统的通过变量)的独立力源. 而且存在有，与环节自身上的力无关，其速度差由系统中其它环节的变量(速度差、力)规定的从属速度源，以及力的大小由其它环节的速度差或力规定的从属力源.

表 3

系统	独立横断变量源	独立通过变量源	从属横断变量源	从属通过变量源
电气系统	 独立电压源	 独立电流源	 从属电压源	 从属电流源
直线运动系统	 独立速度源	 独立力源	 从属速度源	 从属力源
回转运动系统	 独立角速度源	 独立力矩源	 从属角速度源	 从属力矩源
流体系统	 独立压力源	 独立流量源	 从属压力源	 从属流量源

1) 指定横断变量和通过变量后，才能着手环节的分类.

在回转运动系统及流体系统中，也有与其各自的横断变量及通过变量对应的独立和从属角速度源、力矩源、压力源、流量源．以后，这些环节将用表3上的符号表示．为了将一般物理系统统一进行讨论，便于将这些环节按其共同特性分类，给出如下定义．

［定义2］

(i) 可以独立指定横断变量的环节称为独立横断变量源；

(ii) 可以独立指定通过变量的环节称为独立通过变量源；

(iii) 横断变量值由其它环节的变量确定的横断变量源，称为从属横断变量源；

(iv) 通过变量值由其它环节的变量确定的通过变量源，称为从属通过变量源．

此外假定，表3各环节上标的符号 ← 和 +、− 是用来指定各环节用图表示时枝的方向．例如，图14的独立力源表示作用在端子 ⓐ 上外力为 $\hat{F}$ 的力源．显然，对于回转运动系统和流体系统中的环节，也可以由直线运动系统和电气系统类推．

图14　独立力源

2 端子对环节[45]

其次我们来证明，若把以前讲的1端子对环节和从属源组合起来，则可以等效地表示线性物理系统中经常遇到杠杆（直线运动系统）、齿轮（回转运动系统）、两个流出流入口的贮水箱（流体系统）、马达（电气－回转系统）等典型2端子对环节．

首先看一下直线运动系统的理想杠杆（图15(a)）．所谓理想杠杆，就是将实际杠杆的质量和摩擦视为零，把端子对之间的特性线性化后的理想环节，其端子特性可以用

$$\begin{bmatrix} F_1(t) \\ v_2(t) \end{bmatrix} = \begin{bmatrix} 0 & n \\ -n & 0 \end{bmatrix} \begin{bmatrix} v_1(t) \\ F_2(t) \end{bmatrix}, \quad n \triangleq \frac{l_2}{l_1} \tag{12}$$

表示．式中 v_1、v_2 分别为以适当的固定点为基准的端子 ⓐ、ⓑ 的速度差．象电气系统中的理想变压器可以用从属电压源和从属电流源等效表示一样，满足(12)式端子特性的理想杠杆的等效模型也可以用从属速度源和从属力源组成（图15(b)）．

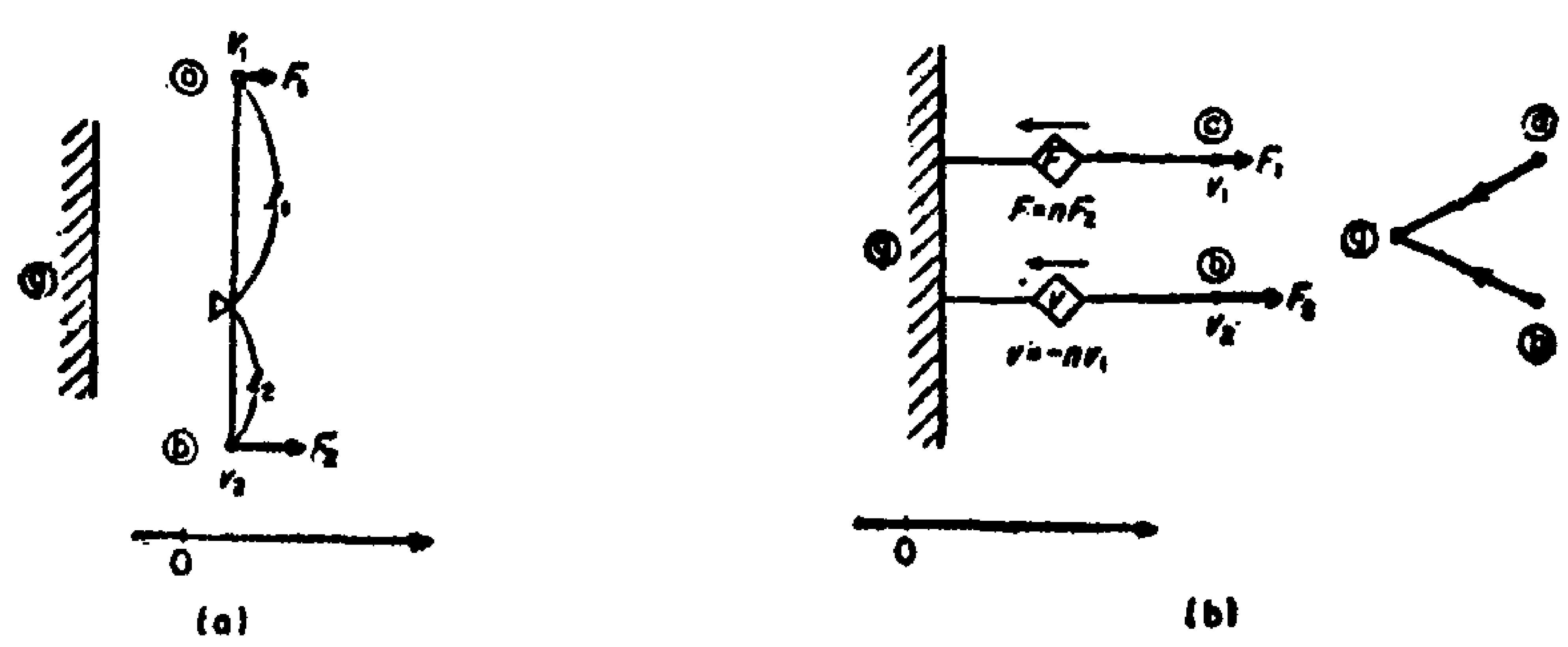


图15　理想杠杆及其等效系统

考虑到杠杆的质量和摩擦时，更实际的杠杆端子特性，在所考虑的线性范围内，可以用

$$\begin{bmatrix} F_1(t) \\ v_2(t) \end{bmatrix} = \begin{bmatrix} B+M\dfrac{d}{dt} & n \\ -n & 0 \end{bmatrix} \begin{bmatrix} v_1(t) \\ F_2(t) \end{bmatrix} \tag{13}$$

表示．式中 B、M 分别为考虑到支点摩擦及杠杆质量后引入的常数．由上面的端子特性可见，考虑到质量和摩擦的杠杆，可以用理想杠杆、刚体及缓冲器的组合表示(图 16(a))．因此，杠杆的有效模型如图 16(b)所示．

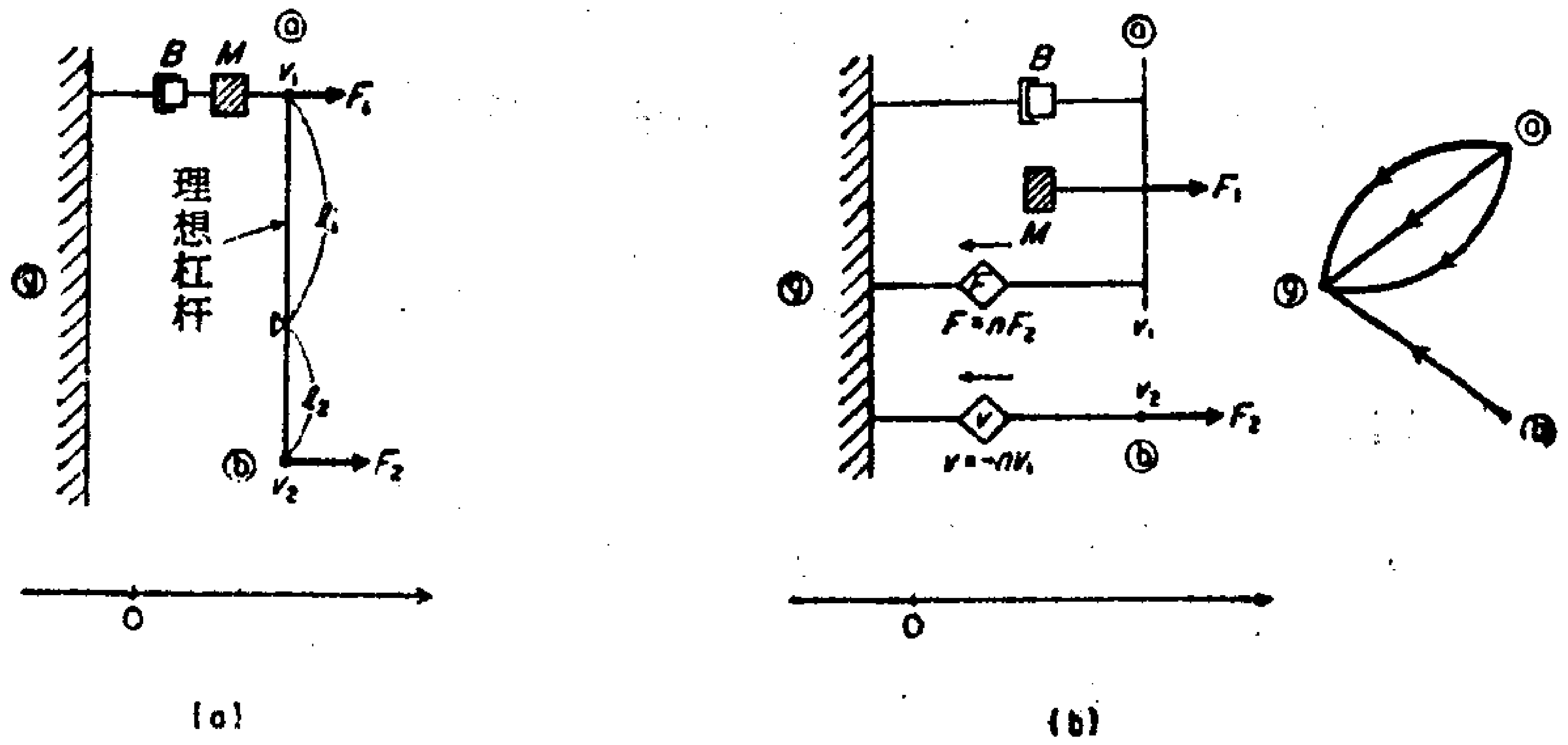


图 16　考虑质量和摩擦的杠杆

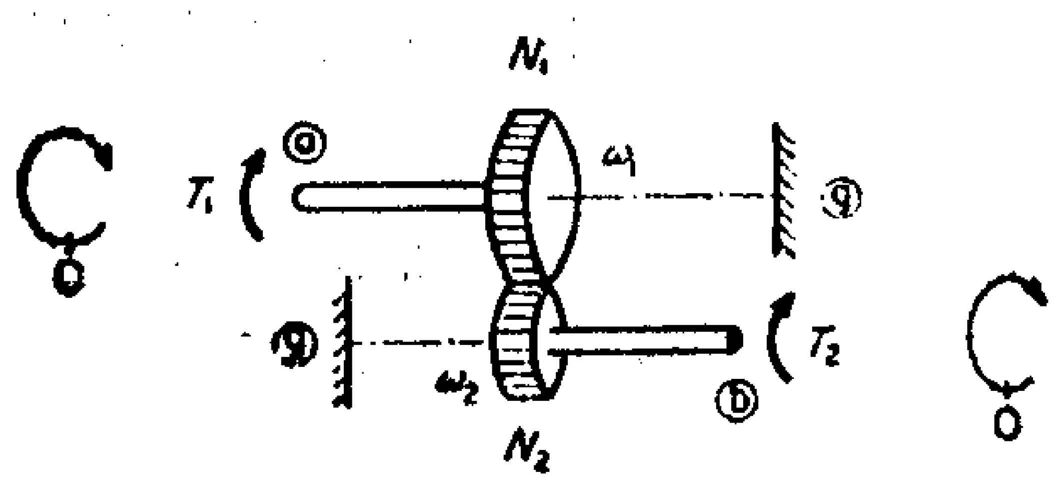


图 17　齿轮

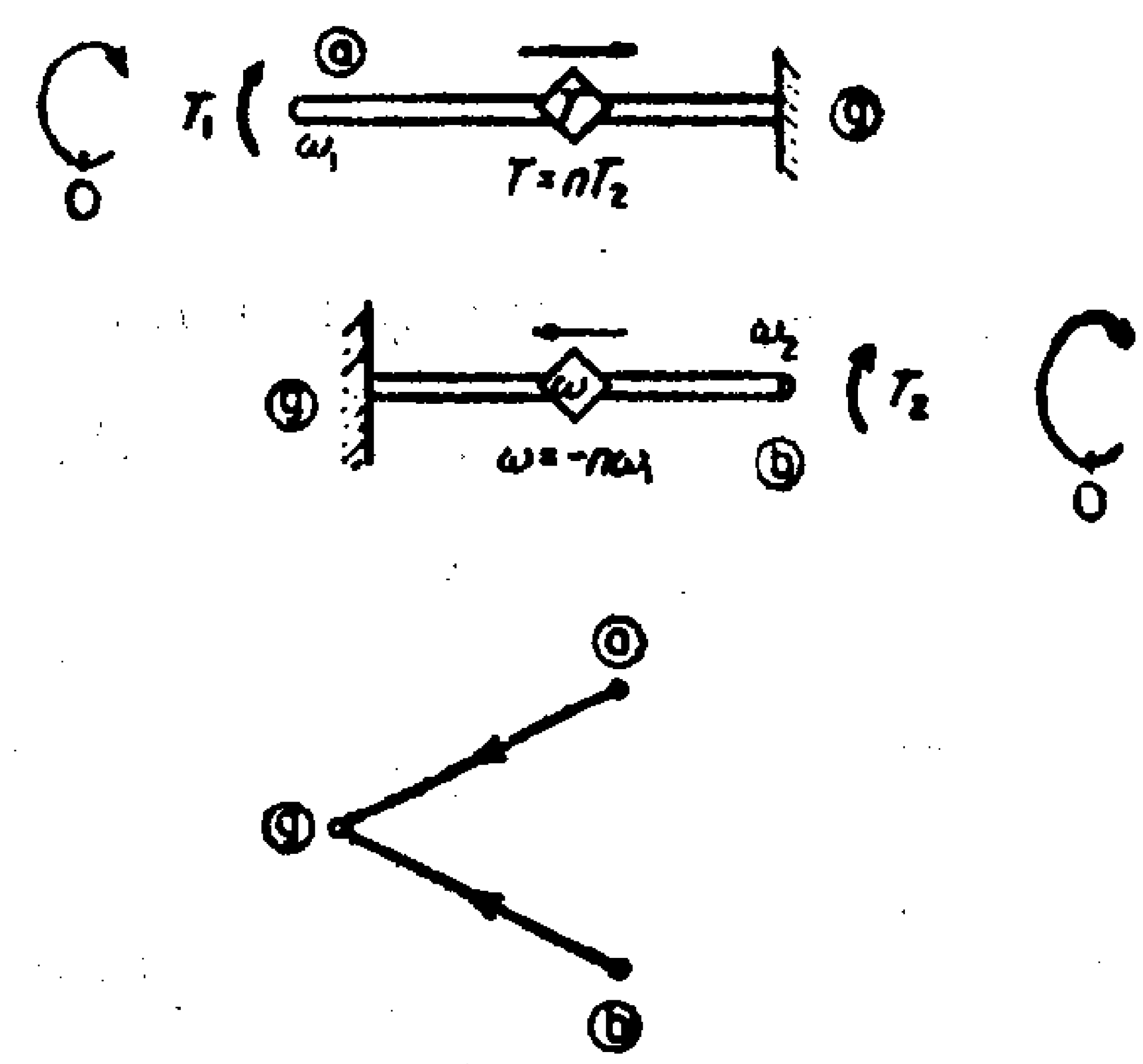


图 18　理想齿轮的等效模型

回转运动系统中齿轮(图 17)和上面的杠杆一样．忽略惯性和摩擦的理想齿轮，其端子特性可以用

$$\begin{bmatrix} T_1(t) \\ \omega_2(t) \end{bmatrix} = \begin{bmatrix} 0 & n \\ -n & 0 \end{bmatrix} \begin{bmatrix} \omega_1(t) \\ T_2(t) \end{bmatrix}, \quad n \triangleq \frac{N_1}{N_2} \tag{14}$$

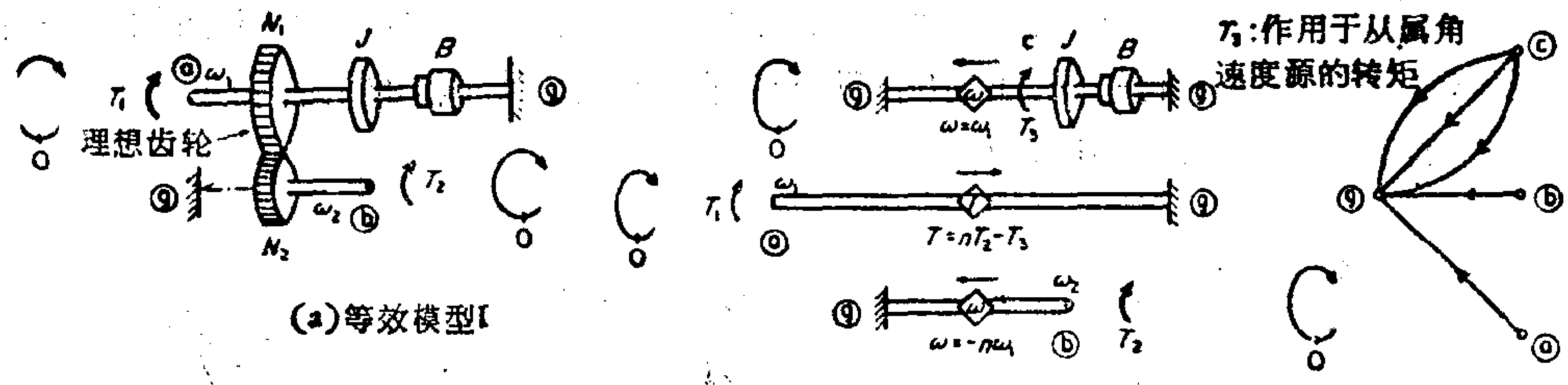
表示，其等效模型如图 18 所示．图 19 是考虑到惯性和摩擦时齿轮的等效模型，其端子特性可以用

$$\begin{bmatrix} T_1(t) \\ \omega_2(t) \end{bmatrix} = \begin{bmatrix} B+J\dfrac{d}{dt} & n \\ -n & 0 \end{bmatrix} \begin{bmatrix} \omega_1(t) \\ T_2(t) \end{bmatrix} \tag{15}$$

表示，式中 ω_1、ω_2 各为以适当的固定点为基准的角速度差．

具有两个流入流出口的贮水箱(图 20)，因为当可以忽略流体惯性时，其端子特性可以用

$$\begin{bmatrix} P_1(t) \\ f_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & C\dfrac{d}{dt} \end{bmatrix} \begin{bmatrix} f_1(t) \\ P_2(t) \end{bmatrix} \quad C: \text{常数} \tag{16}$$



(a)等效模型I

(b)等效模型II

图 19 考虑到惯性及摩擦时齿轮的等效模型

而考虑流体惯性时，可以用

$$\begin{bmatrix} P_1(t) \\ f_2(t) \end{bmatrix} = \begin{bmatrix} 2R_{12} - R_{12}^2 C\dfrac{d}{dt} & 1 - R_{12}C\dfrac{d}{dt} \\ R_{12}C\dfrac{d}{dt} - 1 & C\dfrac{d}{dt} \end{bmatrix} \begin{bmatrix} f_1(t) \\ P_2(t) \end{bmatrix}$$

$$= \begin{bmatrix} 1 & -R_{12} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & C\dfrac{d}{dt} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ R_{12} & 1 \end{bmatrix} \begin{bmatrix} f_1(t) \\ P_2(t) \end{bmatrix}$$

$$C,\ R_{12}: \text{常数} \tag{17}$$

表示，则两种情况下的等效模型如图 21(a)、(b)所示. (17)式中的 R_{12} 是考虑惯性时引入的常数.

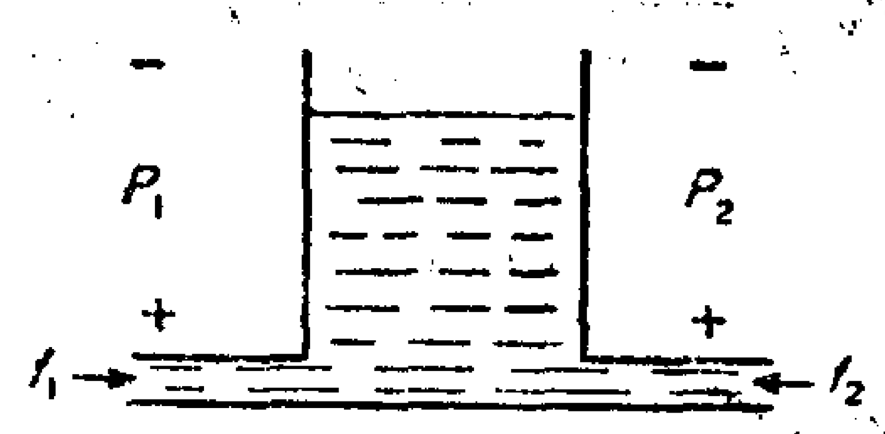
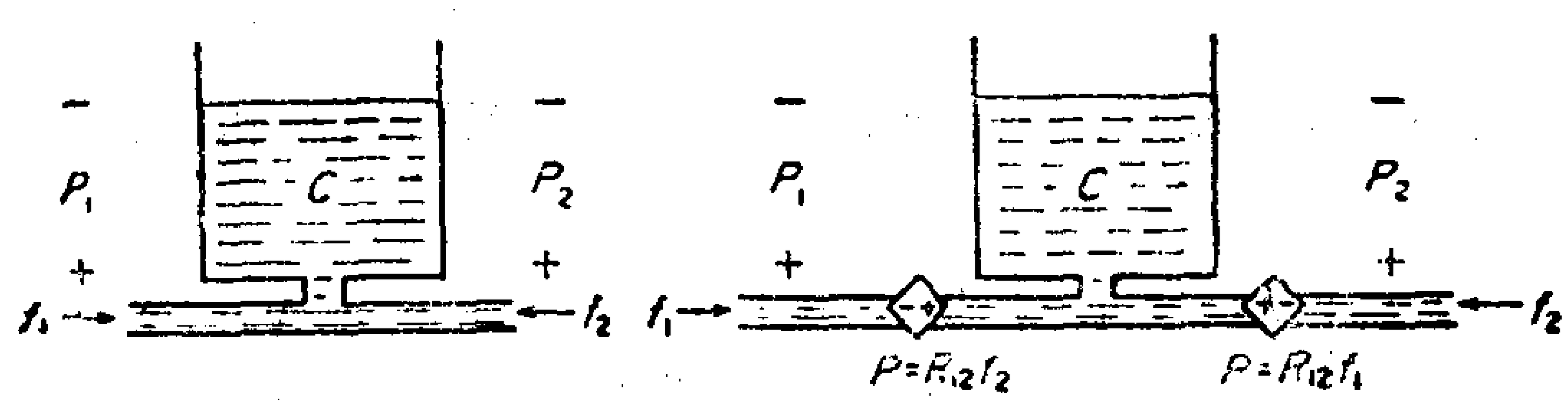


图 20 具有两个流出流入口的贮水箱

上面三个例子，杠杆、齿轮及两个流出流入口的贮水箱，都分别是属于直线运动系统、回转运动系统及流体系统这样一些单一系统中的环节. 作为跨越几个不同系统的例子，下面我们来证明，2 端子对环节马达(图 22)也可以用基本环节等效表示. 马达是电气-回转运动系统中具有代表性的环节. 现设励磁电流一定、马达电枢侧端子对电压、电流分别为 v_1、i_1，作用在转子上的力矩为 T_2，以适



(a)忽略流体惯性时

(b)考虑流体惯性时

图 21 具有两个流出流入口的贮水箱及其等效模型

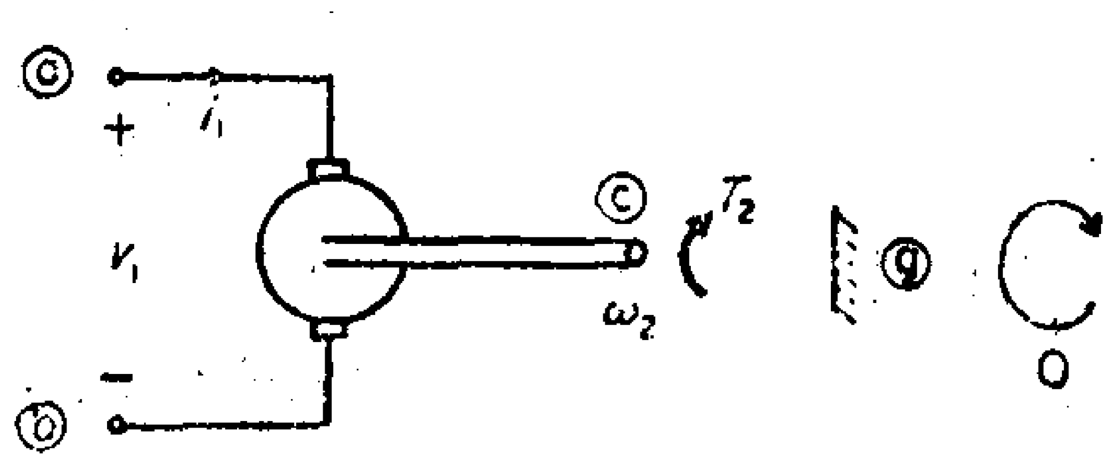


图 22 马达

当的固定点为基准的转子角速度差为 ω_2. 这时，若电枢的电阻和电感及转子的质量和摩擦都小到可以忽略不计的程度，则线性马达的端子特性为

$$\begin{bmatrix} v_1(t) \\ T_2(t) \end{bmatrix} = \begin{bmatrix} 0 & K_m \\ -K_m & 0 \end{bmatrix} \begin{bmatrix} i_1(t) \\ \omega_2(t) \end{bmatrix} \tag{18}$$

而考虑到上述诸因素时，可用

$$\begin{bmatrix} v_1(t) \\ T_2(t) \end{bmatrix} = \begin{bmatrix} R_m + L_m \dfrac{d}{dt} & K_m \\ -K_m & b_m + j_m \dfrac{d}{dt} \end{bmatrix} \begin{bmatrix} i_1(t) \\ \omega_2(t) \end{bmatrix} \tag{19}$$

表示. 式中 K_m 是由励磁电流决定的常数，R_m、L_m 是电枢的电阻及电感，b_m, j_m 分别表示转子的摩擦和惯性. 显然，端子特性分别用(18)、(19)式表示的两种情况下的等效模型如图 23(a)、(b)所示.

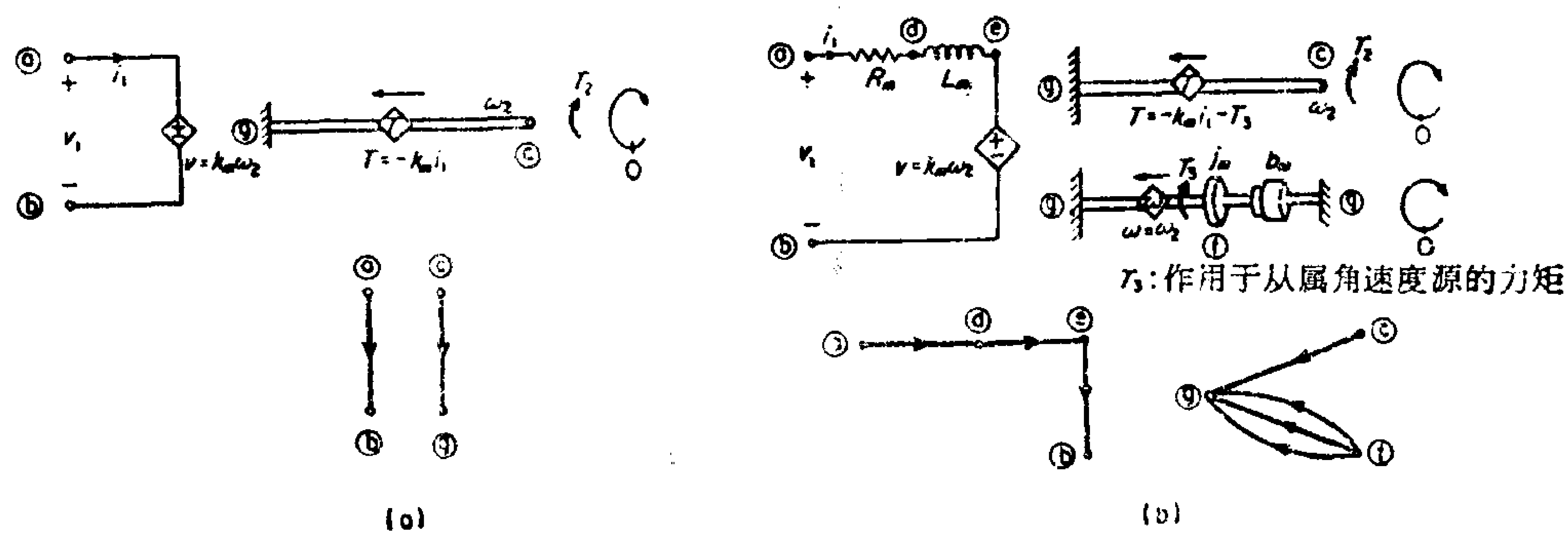


图 23 马达的等效模型及其图

状态方程式的推导方法

由前面讲过的内容显而易见，A-I-2 电气系统状态方程式的推导方法，若将其中的电压换成横断变量、电流换成通过变量、电阻换成代数环节、电容换成一阶横断环节、电感换成一阶通过环节，则该方法也适用于电气系统以外的其它物理系统. 因此，为了避免再重复叙述前述的方法，以下只用实例来说明，该方法也可以用于直线运动系统、回转运动系统、流体系统及电气-回转系统.

(例 1) 直线运动系统的例子

现来讨论图 24 所示系统. 设输入为独立力源的力 $\hat{F}$, 输出为作用于刚体 3 上的力，

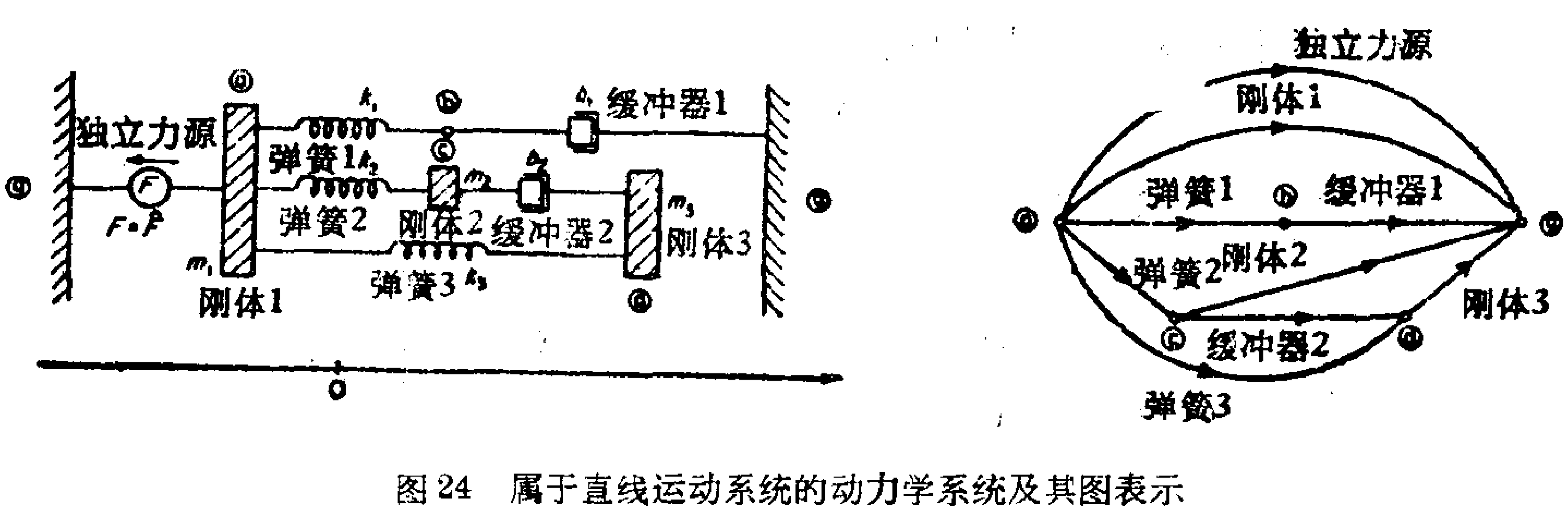


图 24 属于直线运动系统的动力学系统及其图表示

电气系统中定义的真树(P 树)(参看 A-I-2),在该系统中变成下列定义内容.

[定义 3]　包含所有横断变量源和所有一阶横断环节，而不含有通过变量源及一阶通过环节的树,叫做真树(P 树).

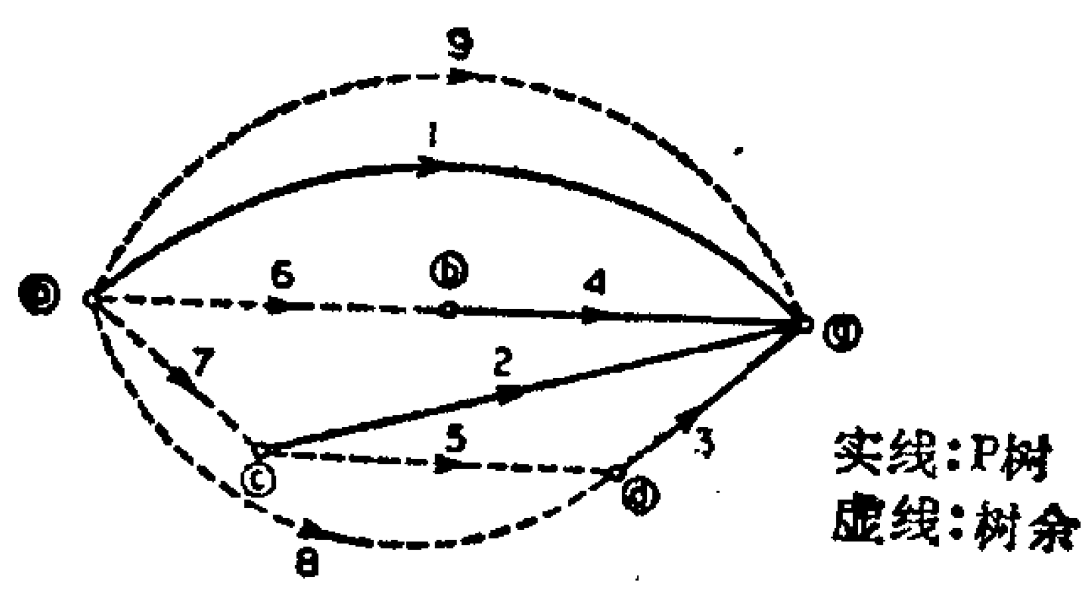


图 25　图 24 上图的 P 树及枝的编号

找出图 24 图的 P 树，按照刚体枝(当然包含在 P 树中)、包含在 P 树中的缓冲器枝、包含在树余中的缓冲器枝、弹簧枝(包含在树余中)及独立力源枝(包含在树余中)的顺序对枝进行编号(参看图 25). 使用该编号及 A-I-2 用过的符号，将该系统的变量(速度差、力)分类如下:

$$\boldsymbol{V}=\begin{bmatrix}\boldsymbol{V}_C\\ \hdashline \boldsymbol{V}_G\\ \hdashline \boldsymbol{V}_R\\ \hdashline \boldsymbol{V}_L\\ \hdashline \boldsymbol{V}_I\end{bmatrix}=\begin{bmatrix}v_1\\ v_2\\ v_3\\ \hdashline v_4\\ \hdashline v_5\\ \hdashline v_6\\ v_7\\ v_8\\ \hdashline v_9\end{bmatrix},$$

$$\boldsymbol{I}=\begin{bmatrix}\boldsymbol{I}_C\\ \hdashline \boldsymbol{I}_G\\ \hdashline \boldsymbol{I}_R\\ \hdashline \boldsymbol{I}_L\\ \hdashline \boldsymbol{I}_I\end{bmatrix}=\begin{bmatrix}F_1\\ F_2\\ F_3\\ \hdashline F_4\\ \hdashline F_5\\ \hdashline F_6\\ F_7\\ F_8\\ \hdashline F_9\end{bmatrix}$$

从上式可见，由电气系统的状态变量类推，可以令 v_1、v_2、v_3、F_6、F_7、F_8 为该系统的状态变量，即可设刚体的速度差及作用于弹簧上的力为该系统(属于直线运动系统)的状态变量. 那末,关于 P 树的基本割集矩阵为

$$\begin{array}{cc} & \overbrace{\begin{array}{ccccccccc}1&2&3&4&5&6&7&8&9\end{array}}^{\text{枝}} \\ \boldsymbol{Q}_f= & \left[\begin{array}{cccc:ccccc}1&0&0&0&0&1&1&1&1\\0&1&0&0&1&0&-1&0&0\\0&0&1&0&-1&0&0&-1&0\\0&0&0&1&0&-1&0&0&0\end{array}\right]=[\boldsymbol{I}:\boldsymbol{F}]\end{array}$$

如 A–I–2 那样，将上式中的部分矩阵 $\boldsymbol{F}$ 分解为下列分块矩阵

$$\boldsymbol{F}=\left[\begin{array}{c:ccc:c} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & -1 & 0 \\ \hdashline 0 & -1 & 0 & 0 & 0 \end{array}\right]=\left[\begin{array}{c:c:c} \boldsymbol{F}_{CR} & \boldsymbol{F}_{CL} & \boldsymbol{F}_{CI} \\ \hdashline \boldsymbol{F}_{GR} & \boldsymbol{F}_{GL} & \boldsymbol{F}_{GI} \end{array}\right] \tag{20}$$

利用 $\boldsymbol{Q}_f(\boldsymbol{F})$ 求出基本割集方程式、基本回路方程式如下：

$$\left[\begin{array}{c} F_1(t) \\ F_2(t) \\ F_3(t) \\ \hdashline F_4(t) \end{array}\right]=\left[\begin{array}{ccccc} 0 & -1 & -1 & -1 & -1 \\ -1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{array}\right]\left[\begin{array}{c} F_5(t) \\ \hdashline F_6(t) \\ F_7(t) \\ F_8(t) \\ \hdashline F_9(t) \end{array}\right] \tag{21}$$

$$\left[\begin{array}{c} v_5(t) \\ \hdashline v_6(t) \\ v_7(t) \\ v_8(t) \\ \hdashline v_9(t) \end{array}\right]=\left[\begin{array}{cccc} 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right]\left[\begin{array}{c} v_1(t) \\ v_2(t) \\ v_3(t) \\ \hdashline v_4(t) \end{array}\right] \tag{22}$$

此外，独立力源、缓冲器、刚体、弹簧的端子特性式可用下列诸式给出：

$$F_9(t)=\hat{F}(t) \tag{23}$$

$$\left[\begin{array}{c} F_4(t) \\ \hdashline v_5(t) \end{array}\right]=\left[\begin{array}{cc} b_1 & 0 \\ 0 & b_2^{-1} \end{array}\right]\left[\begin{array}{c} v_4(t) \\ \hdashline F_5(t) \end{array}\right]$$

$$\begin{bmatrix} F_1(t) \\ F_2(t) \\ F_3(t) \end{bmatrix}=\begin{bmatrix} m_1 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_3 \end{bmatrix}\begin{bmatrix} \dot{v}_1(t) \\ \dot{v}_2(t) \\ \dot{v}_3(t) \end{bmatrix}$$

$$\begin{bmatrix} v_6(t) \\ v_7(t) \\ v_8(t) \end{bmatrix}=\begin{bmatrix} k^{-1} & 0 & 0 \\ 0 & k_2^{-1} & 0 \\ 0 & 0 & k_3^{-1} \end{bmatrix}\begin{bmatrix} \dot{F}_6(t) \\ \dot{F}_7(t) \\ \dot{F}_8(t) \end{bmatrix}$$

因此，若使用 A–I–2 用过的符号，则

$$\left[\begin{array}{c:c} b_1 & 0 \\ \hdashline 0 & b_2^{-1} \end{array}\right]=\left[\begin{array}{c:c} \boldsymbol{G}_G & \boldsymbol{0} \\ \hdashline \boldsymbol{0} & \boldsymbol{R}_R \end{array}\right] \tag{24}$$

$$\begin{bmatrix} m_1 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_3 \end{bmatrix}=\boldsymbol{C}_C \tag{25}$$

$$\begin{bmatrix} k_1^{-1} & 0 & 0 \\ 0 & k_2^{-1} & 0 \\ 0 & 0 & k_3^{-1} \end{bmatrix}=\boldsymbol{L}_{LL} \tag{26}$$

由(20)，(24)—(26)式可以得出中间标准形(A-I-2 中的(57)式)如下

$$
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & m_1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & m_2 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & m_3 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & k_1^{-1} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & k_2^{-1} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & k_3^{-1}
\end{bmatrix}
\begin{bmatrix}
v_4(t) \\ F_5(t) \\ \dot{v}_1(t) \\ \dot{v}_2(t) \\ \dot{v}_3(t) \\ \dot{F}_6(t) \\ \dot{F}_7(t) \\ \dot{F}_8(t)
\end{bmatrix}
$$

$$
=\begin{bmatrix}
0 & 0 & 0 & b_1^{-1} & 0 & 0 \\
0 & b_2 & -b_2 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & -1 & -1 \\
0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 & 0 \\
1 & -1 & 0 & 0 & 0 & 0 \\
1 & 0 & -1 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
v_1(t) \\ v_2(t) \\ v_3(t) \\ F_6(t) \\ F_7(t) \\ F_8(t)
\end{bmatrix}
+\begin{bmatrix}
0 \\ 0 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0
\end{bmatrix} F_9(t) \tag{27}
$$

该式左边系数矩阵是三角形矩阵，且其对角线元素不全为零，则是正则的．因此，在(27)式两边左乘以该系数矩阵的逆矩阵，或者对(27)式进行适当的基本行变换，使左边的系数矩阵变成单位矩阵，则得

$$
\begin{bmatrix}
v_4(t) \\ F_5(t) \\ \dot{v}_1(t) \\ \dot{v}_2(t) \\ \dot{v}_3(t) \\ \dot{F}_6(t) \\ \dot{F}_7(t) \\ \dot{F}_8(t)
\end{bmatrix}
=\begin{bmatrix}
0 & 0 & 0 & b^{-1} & 0 & 0 \\
0 & b_2 & -b_2 & 0 & 0 & 0 \\
0 & 0 & 0 & -m_1^{-1} & -m_1^{-1} & -m_1^{-1} \\
0 & -m_2^{-1}b_2 & m_2^{-1}b_2 & 0 & m_2^{-1} & 0 \\
0 & m_3^{-1}b_2 & -m_3^{-1}b_2 & 0 & 0 & m_3^{-1} \\
k_1 & 0 & 0 & -k_1b_1^{-1} & 0 & 0 \\
k_2 & -k_2 & 0 & 0 & 0 & 0 \\
k_3 & 0 & -k_3 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
v_1(t) \\ v_2(t) \\ v_3(t) \\ F_6(t) \\ F_7(t) \\ F_8(t)
\end{bmatrix}
$$

$$
+\begin{bmatrix}
0 \\ 0 \\ -m_1^{-1} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0
\end{bmatrix} F_9(t) \tag{28}
$$

将独立力源的端子特性式(23)代入该式，取下侧 6 行便得到状态方程式

$$
\begin{bmatrix} \dot{v}_1(t) \\ \dot{v}_2(t) \\ \dot{v}_3(t) \\ \dot{F}_6(t) \\ \dot{F}_7(t) \\ \dot{F}_8(t) \end{bmatrix}
=
\begin{bmatrix}
0 & 0 & 0 & -m_1^{-1} & -m_1^{-1} & -m_1^{-1} \\
0 & -m_2^{-1}b_2 & m_2^{-1}b_2 & 0 & m_2^{-1} & 0 \\
0 & m_3^{-1}b_2 & -m_3^{-1}b_2 & 0 & 0 & m_3^{-1} \\
k_1 & 0 & 0 & -k_1 b_1^{-1} & 0 & 0 \\
k_2 & -k_2 & 0 & 0 & 0 & 0 \\
k_3 & 0 & -k_3 & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix} v_1(t) \\ v_2(t) \\ v_3(t) \\ F_6(t) \\ F_7(t) \\ F_8(t) \end{bmatrix}
+
\begin{bmatrix} -m_1^{-1} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \hat{F}(t)
\tag{29}
$$

此外，因输出 y 是 F_3，根据(21)式第 3 式则可表示成

$$y(t)=F_5(t)+F_8(t)$$

将(28)式中的第 2 式代入该式中的 $F_5(t)$，便得到下列输出方程式

$$
y(t)=\begin{bmatrix} 0 & b_2 & -b_2 & 0 & 0 & 1 \end{bmatrix}
\begin{bmatrix} v_1(t) \\ v_2(t) \\ v_3(t) \\ F_6(t) \\ F_7(t) \\ F_8(t) \end{bmatrix}.
\tag{30}
$$

(例 2)　回转运动系统的例子.

现在我们来推导图 26 所示属于回转运动系统的动力学系统的状态方程式. 设输入是独立角速度源 ⓓ 的角速度差，输出是作用于端子 ⓓ 独立角速度源上的力矩. 首先，利用理想齿轮的等效模型(图 18)，将图 26 所示系统表示在图 27.

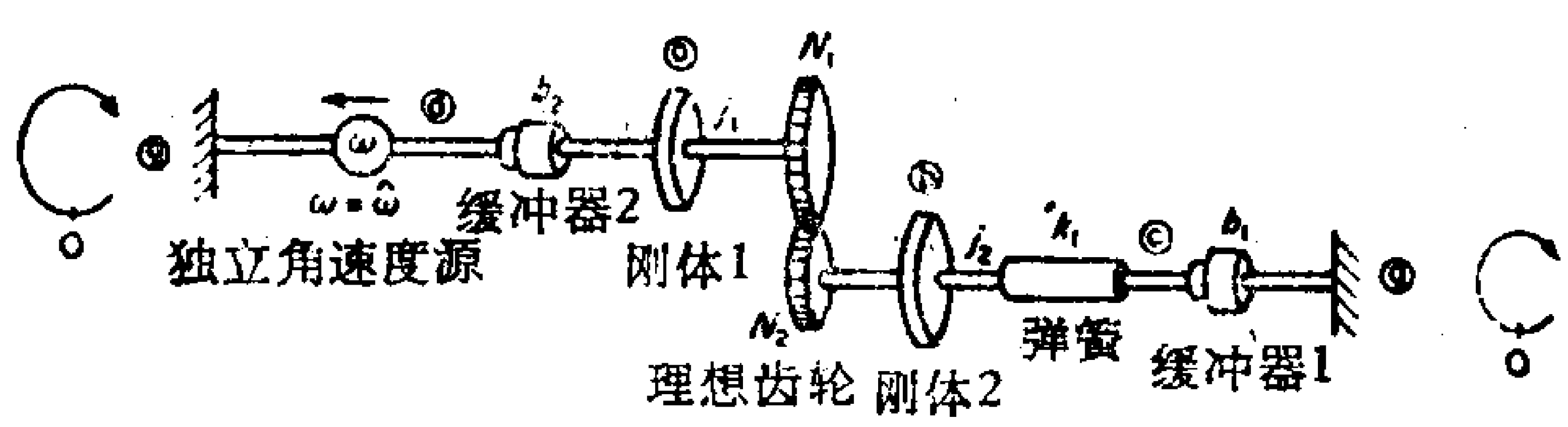


图 26　属于回转运动系统的动力学系统

这里，将电气系统中定义过的正规树(N 树)(参看 A-I-2)重新定义如下.

[定义 4]　包含所有独立及从属横断变量源及尽可能多的一阶横断环节，包含所有独立及从属通过变量源和尽可能多的一阶通过环节的树，叫做正规树(N 树).

找出图 27 上图的 N 树，按照从属角速度源枝、独立角速度源枝、包含在 N 树中的刚体枝、包含在 N 树中的缓冲器枝、包含在树余中的刚体枝、包含在树余中的缓冲器枝、(包含在树余中的)弹簧枝、从属力矩源枝的顺序，对枝编号后示于图 28. 使用该编号和

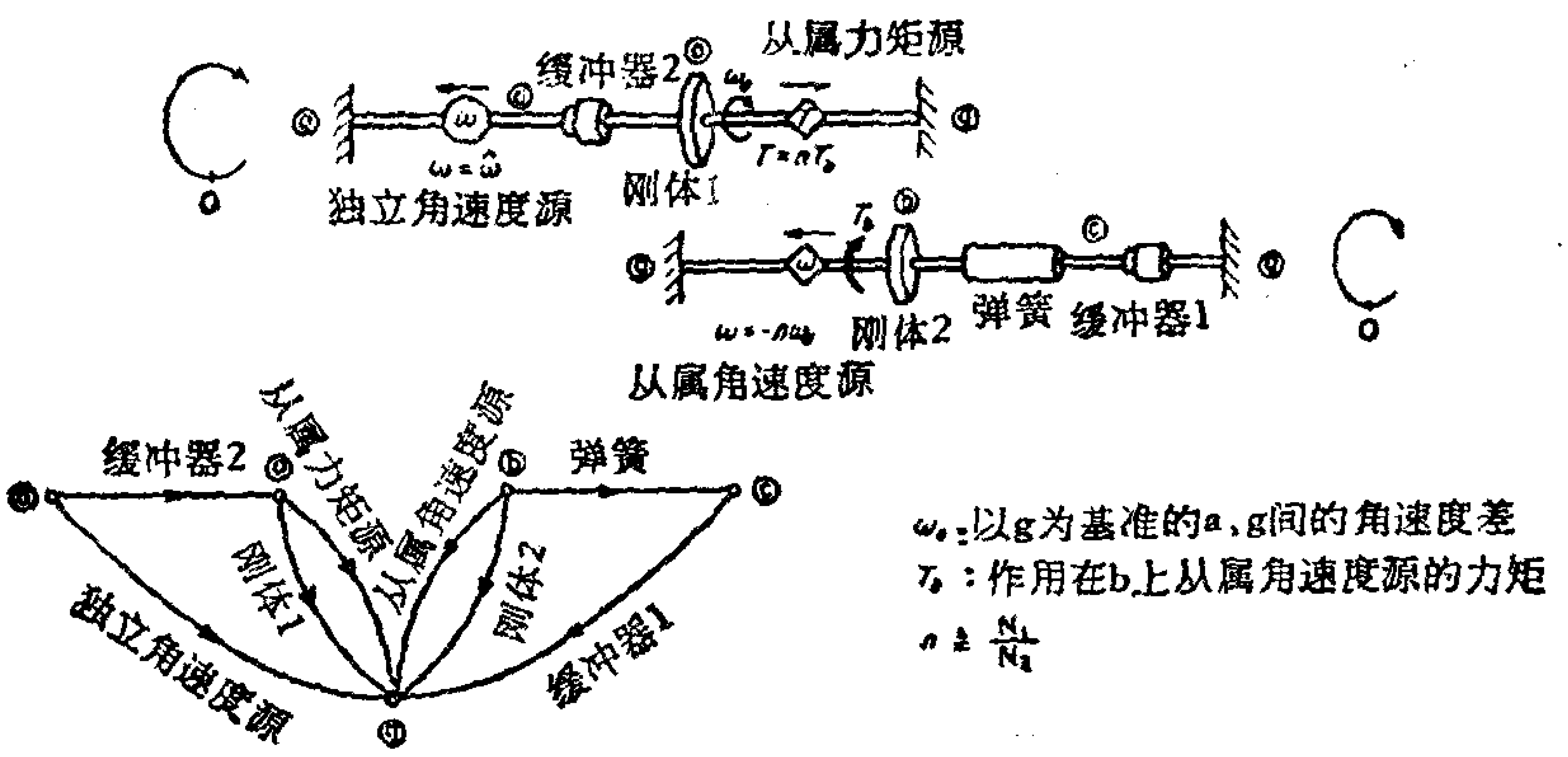


图 27　图 26 所示系统的等效模型及其图表示

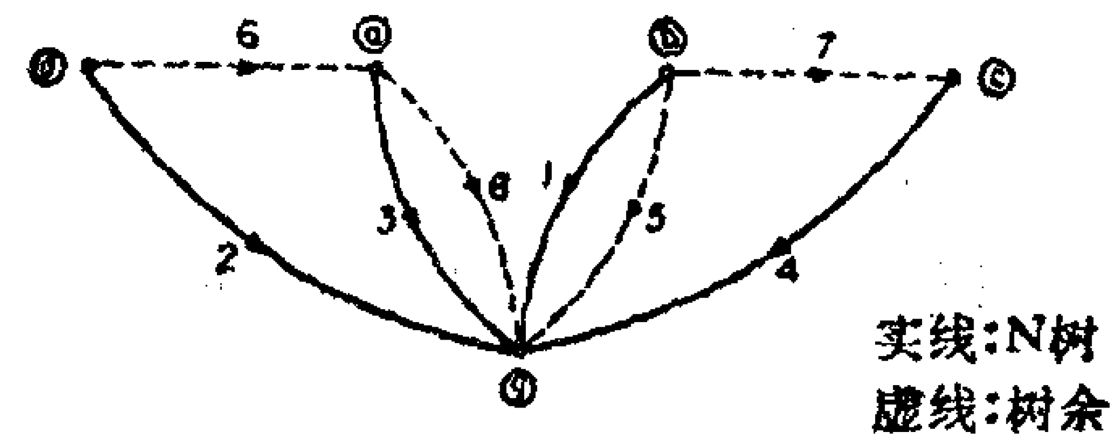


图 28　图 27 所示图中 N 树枝的编号

A-I-2 用过的符号,将各枝的变量(角速度差、力矩)分类如下:

$$
\boldsymbol{V}=\begin{bmatrix} \boldsymbol{V}_A \\ \boldsymbol{V}_V \\ \boldsymbol{V}_C \\ \boldsymbol{V}_G \\ \boldsymbol{V}_S \\ \boldsymbol{V}_R \\ \boldsymbol{V}_L \\ \boldsymbol{V}_B \end{bmatrix}=\begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \\ \omega_4 \\ \omega_5 \\ \omega_6 \\ \omega_7 \\ \omega_8 \end{bmatrix}, \qquad \boldsymbol{I}=\begin{bmatrix} \boldsymbol{I}_A \\ \boldsymbol{I}_V \\ \boldsymbol{I}_C \\ \boldsymbol{I}_G \\ \boldsymbol{I}_S \\ \boldsymbol{I}_R \\ \boldsymbol{I}_L \\ \boldsymbol{I}_B \end{bmatrix}=\begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \\ T_5 \\ T_6 \\ T_7 \\ T_8 \end{bmatrix}
$$

由上面分类可见,该系统的补充状态变量是 ω_3、ω_5、T_7. 亦即,属于回转运动系统的动力学系统,其补充状态变量是刚体的角速度差和作用在弹簧上的力矩.

那么,关于 N 树的基本割集矩阵是

$$
\begin{array}{c}
\overbrace{\begin{array}{cccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array}}^{\text{枝}} \\
\boldsymbol{Q}_f=\left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 \end{array}\right]=[\boldsymbol{I} \mid \boldsymbol{F}]
\end{array}
$$

使用 A-I-2 用过的符号,则式中部分矩阵 $\boldsymbol{F}$ 可以分解成如下分块矩阵

$$\boldsymbol{F}=\left[\begin{array}{c:c:c:c} 1 & 0 & 1 & 0 \\ \hdashline 0 & 1 & 0 & 0 \\ \hdashline 0 & -1 & 0 & 1 \\ \hdashline 0 & 0 & -1 & 0 \end{array}\right]=\left[\begin{array}{c:c:c:c} \boldsymbol{F}_{AS} & \boldsymbol{F}_{AR} & \boldsymbol{F}_{AL} & \boldsymbol{F}_{AB} \\ \hdashline \boldsymbol{F}_{VS} & \boldsymbol{F}_{VR} & \boldsymbol{F}_{VL} & \boldsymbol{F}_{VB} \\ \hdashline \boldsymbol{F}_{CS} & \boldsymbol{F}_{CR} & \boldsymbol{F}_{CL} & \boldsymbol{F}_{CB} \\ \hdashline \boldsymbol{0} & \boldsymbol{F}_{GR} & \boldsymbol{F}_{GL} & \boldsymbol{F}_{GB} \end{array}\right] \tag{31}$$

基本割集方程式和基本回路方程式分别求出如下

$$\left[\begin{array}{c} T_1(t) \\ \hdashline T_2(t) \\ \hdashline T_3(t) \\ \hdashline T_4(t) \end{array}\right]=\begin{bmatrix} -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}\left[\begin{array}{c} T_5(t) \\ \hdashline T_6(t) \\ \hdashline T_7(t) \\ \hdashline T_8(t) \end{array}\right] \tag{32}$$

$$\left[\begin{array}{c} \omega_5(t) \\ \hdashline \omega_6(t) \\ \hdashline \omega_7(t) \\ \hdashline \omega_8(t) \end{array}\right]=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}\left[\begin{array}{c} \omega_1(t) \\ \hdashline \omega_2(t) \\ \hdashline \omega_3(t) \\ \hdashline \omega_4(t) \end{array}\right] \tag{33}$$

此外，独立角速度源.从属角速度源及从属力矩源、缓冲器、刚体、弹簧的端子特性式可分别给出如下

$$\omega_2(t)=\hat{\omega}(t) \tag{34}$$

$$\left[\begin{array}{c} \omega_1(t) \\ \hdashline T_8(t) \end{array}\right]=\begin{bmatrix} 0 & -n \\ n & 0 \end{bmatrix}\left[\begin{array}{c} T_1(t) \\ \hdashline \omega_8(t) \end{array}\right] \tag{35}$$

$$\left[\begin{array}{c} T_4(t) \\ \hdashline \omega_6(t) \end{array}\right]=\begin{bmatrix} b_1 & 0 \\ 0 & b_2^{-1} \end{bmatrix}\left[\begin{array}{c} \omega_4(t) \\ \hdashline T_6(t) \end{array}\right] \tag{36}$$

$$\left[\begin{array}{c} T_3(t) \\ \hdashline T_5(t) \end{array}\right]=\begin{bmatrix} j_1 & 0 \\ 0 & j_2 \end{bmatrix}\left[\begin{array}{c} \dot{\omega}_3(t) \\ \hdashline \dot{\omega}_5(t) \end{array}\right] \tag{37}$$

$$\omega_7(t)=k_1^{-1}\dot{T}_7(t) \tag{38}$$

利用(32)、(33)式，则由(35)式可得出与 A-I-2 中(48)式相当的如下表示式

$$\left[\begin{array}{c} \omega_1(t) \\ \hdashline T_8(t) \end{array}\right]=\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & -n & 0 & 0 \\ 0 & 0 & 0 & 0 & -n & 0 & -n & 0 \end{bmatrix}\left[\begin{array}{c} \omega_1(t) \\ \hdashline T_8(t) \\ \hdashline \omega_4(t) \\ \hdashline T_6(t) \\ \hdashline T_5(t) \\ \hdashline \omega_3(t) \\ \hdashline T_7(t) \\ \hdashline \omega_2(t) \end{array}\right] \tag{39}$$

象 A-I-2 中那样，将上式右边的系数矩阵分解成如下分块矩阵

$$\left[\begin{array}{c:c:c:c:c:c:c:c} 0 & 0 & 0 & 0 & 0 & -n & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & -n & 0 & -n & 0 \end{array}\right]=\left[\begin{array}{c:c:c:c:c:c:c:c} \boldsymbol{P}_{11} & \boldsymbol{P}_{12} & \boldsymbol{P}_{13} & \boldsymbol{P}_{14} & \boldsymbol{P}_{15} & \boldsymbol{P}_{17} & \boldsymbol{P}_{18} & \boldsymbol{P}_{19} \\ \hdashline \boldsymbol{P}_{21} & \boldsymbol{P}_{22} & \boldsymbol{P}_{23} & \boldsymbol{P}_{24} & \boldsymbol{P}_{25} & \boldsymbol{P}_{27} & \boldsymbol{P}_{28} & \boldsymbol{P}_{29} \end{array}\right] \tag{40}$$

而且，(36)、(37)、(38)式右边的系数矩阵也可以分解成如下分块矩阵

$$\left[\begin{array}{c:c} b_1 & 0 \\ \hdashline 0 & b_2^{-1} \end{array}\right]=\left[\begin{array}{c:c} \boldsymbol{G}_G & \boldsymbol{0} \\ \hdashline \boldsymbol{0} & \boldsymbol{R}_R \end{array}\right] \tag{41}$$

$$\left[\begin{array}{c:c} j_1 & 0 \\ \hdashline 0 & j_2 \end{array}\right]=\left[\begin{array}{c:c} \boldsymbol{C}_C & \boldsymbol{0} \\ \hdashline \boldsymbol{0} & \boldsymbol{C}_S \end{array}\right] \tag{42}$$

$$k_1^{-1}=\boldsymbol{L}_{LL} \tag{43}$$

利用以上得出的分块矩阵，即利用(31)、(40)～(43)式，便可得到中间标准形(A-I-2 中的(57)式)

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & nj_2 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & j_1 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & k_1^{-1} \end{bmatrix}\begin{bmatrix} \omega_1(t) \\ T_8(t) \\ \omega_4(t) \\ T_6(t) \\ \dot{\omega}_5(t) \\ \dot{\omega}_3(t) \\ \dot{T}_7(t) \end{bmatrix}$$

$$=\begin{bmatrix} 0 & -n & 0 \\ 0 & 0 & -n \\ 0 & 0 & b_1^{-1} \\ 0 & -b_2 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} \omega_5(t) \\ \omega_3(t) \\ T_7(t) \end{bmatrix}+\begin{bmatrix} 0 \\ 0 \\ 0 \\ b_2 \\ 0 \\ 0 \\ 0 \end{bmatrix}\omega_2(t) \tag{44}$$

因为在该式左边的系数矩阵(A-I-2 中用符号 $\boldsymbol{S}$ 表示的矩阵)中，将第1行加到第5行时，第5行变成零行向量，则该矩阵不是正则的．如 A-I-2 中所述，因此必须按 $\boldsymbol{S}$ 不是正则矩阵的情况处理．在(44)式中，将第1式加到第5式得

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & nj_2 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & j_1 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & k_1^{-1} \end{bmatrix}\begin{bmatrix} \omega_1(t) \\ T_8(t) \\ \omega_4(t) \\ T_6(t) \\ \dot{\omega}_5(t) \\ \dot{\omega}_3(t) \\ \dot{T}_7(t) \end{bmatrix}=\begin{bmatrix} 0 & -n & 0 \\ 0 & 0 & -n \\ 0 & 0 & b_1^{-1} \\ 0 & -b_2 & 0 \\ -1 & -n & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} \omega_5(t) \\ \omega_3(t) \\ T_7(t) \end{bmatrix}+\begin{bmatrix} 0 \\ 0 \\ 0 \\ b_2 \\ 0 \\ 0 \\ 0 \end{bmatrix}\omega_2(t) \tag{45}$$

将该式中第5式改写后变成

$$\omega_5(t) = -n\omega_3(t) \tag{46}$$

(46)式表示刚体1和刚体2的角速度差具有相互依赖关系．因为刚体1和刚体2是通过齿轮连接的，所以这一点从物理意义上也很容易理解．起初我们把 ω_3、ω_5 都作为补充状态变量，由于 ω_3 和 ω_5 相关，故没有必要把 ω_3、ω_5 都作为状态变量．因此，我们从状态变量中去掉 ω_5，采用 ω_3 作为状态变量．将(46)式代入(45)式，消去 ω_5 后，得

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -n^2 j_2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & j_1 & 0 \\ -1 & 0 & 1 & 0 & 0 & k_1^{-1} \end{bmatrix} \begin{bmatrix} \omega_1(t) \\ T_8(t) \\ \omega_4(t) \\ T_6(t) \\ \dot{\omega}_3(t) \\ \dot{T}_7(t) \end{bmatrix} = \begin{bmatrix} -n & 0 \\ 0 & -n \\ 0 & b_1^{-1} \\ -b_2 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \omega_3(t) \\ T_7(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ b_2 \\ 0 \\ 0 \end{bmatrix} \omega_2(t) \tag{47}$$

因该式左边的系数矩阵是正则的，在(47)式两边左乘以该矩阵的逆矩阵，或对(47)式进行适当的基本行变换，使左边系数矩阵变成单位矩阵，则得

$$\begin{bmatrix} \omega_1(t) \\ T_8(t) \\ \omega_4(t) \\ T_6(t) \\ \dot{\omega}_3(t) \\ \dot{T}_7(t) \end{bmatrix} = \begin{bmatrix} -n & 0 \\ -\dfrac{n^2 j_2 b_2}{j_1 + n^2 j_2} & -\dfrac{n j_1}{j_1 + n^2 j_2} \\ 0 & b_1^{-1} \\ -b_2 & 0 \\ -\dfrac{b_2}{j_1 + n^2 j_2} & \dfrac{n}{j_1 + n^2 j_2} \\ -n k_1 & -b_1^{-1} k_1 \end{bmatrix} \begin{bmatrix} \omega_3(t) \\ T_7(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \dfrac{n^2 j_2 b_2}{j_1 + n^2 j_2} \\ 0 \\ b_2 \\ \dfrac{b_2}{j_1 + n^2 j_2} \\ 0 \end{bmatrix} \omega_2(t) \tag{48}$$

将独立角速度源端子特性式(34)代入上式，由下侧2式立即可得到状态方程式

$$\begin{bmatrix} \dot{\omega}_3(t) \\ \dot{T}_7(t) \end{bmatrix} = \begin{bmatrix} -\dfrac{b_2}{j_1 + n^2 j_2} & \dfrac{n}{j_1 + n^2 j_2} \\ -n k_1 & -b_1^{-1} k_1 \end{bmatrix} \begin{bmatrix} \omega_3(t) \\ T_7(t) \end{bmatrix} + \begin{bmatrix} \dfrac{b_2}{j_1 + n^2 j_2} \\ 0 \end{bmatrix} \hat{\omega}(t) \tag{49}$$

此外，因输出 y 是 T_2，则由基本割集方程式得

$$y(t) = -T_6(t)$$

将(48)式中的第4式代入该式，则得输出方程式

$$y(t) = [b_2 \quad 0] \begin{bmatrix} \omega_3(t) \\ T_7(t) \end{bmatrix} - b_2 \hat{\omega}(t) \tag{50}$$

(例 3)　流体系统的例子．

现在，我们再来研究一下 A-I-1 中例2曾经讨论过的流体系统(图29)．在图30中

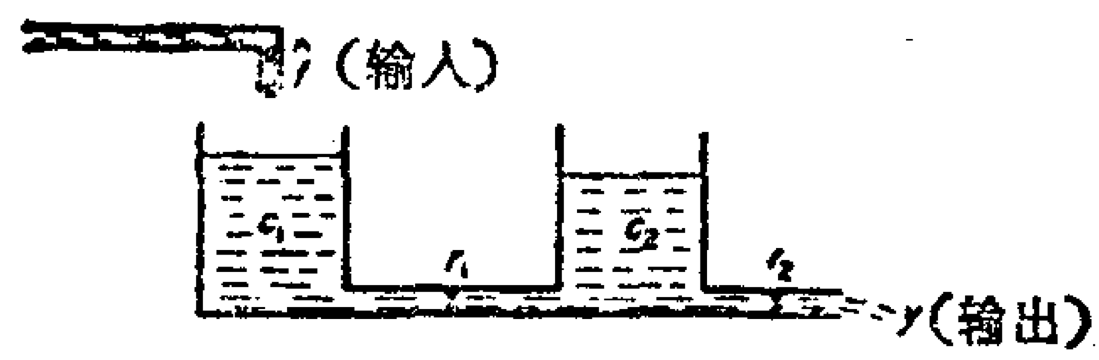


图 29　A-I-1 例 2 中的系统

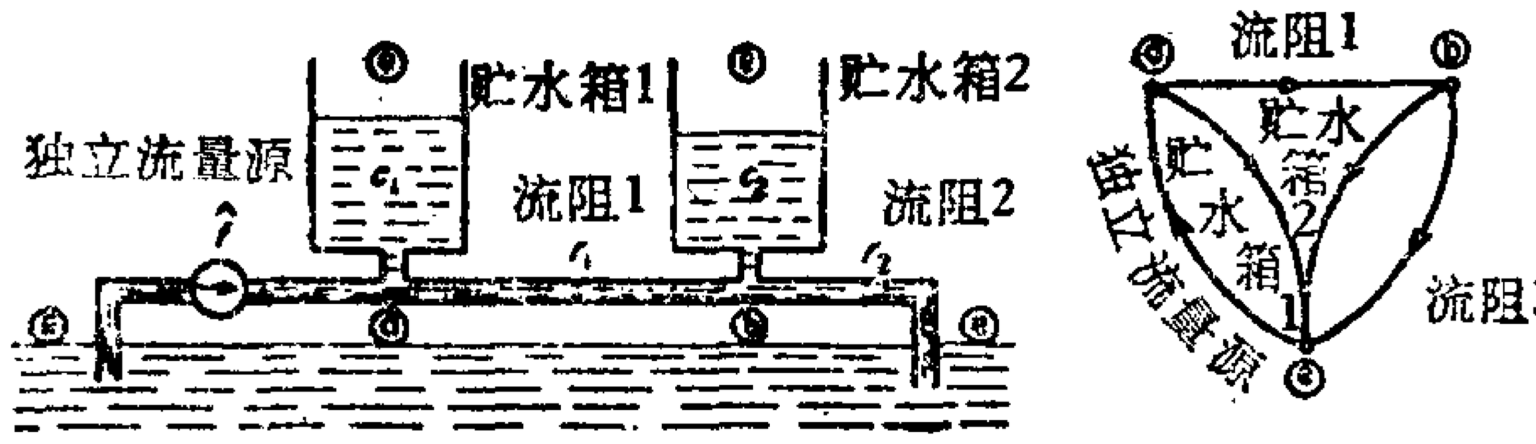


图 30　图 29 所示系统的等效模型及其图表示

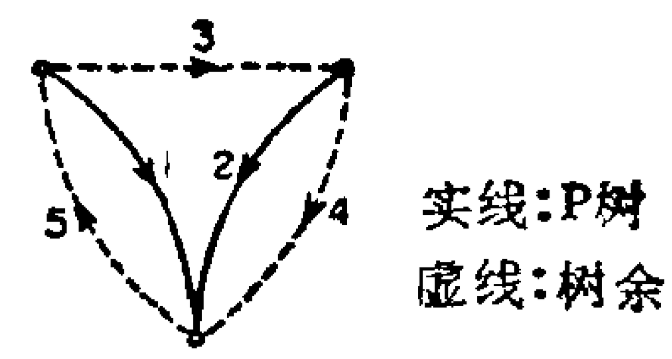


图 31　图 30 所示图的 P 树及枝的编号

将系统用基本环节等效表示. 在该等效模型上，输入是独立流量源的流量，输出是经过流阻 2 的流量.

在图 30 所示图中选取 P 树，并按照贮水箱枝、流阻枝、独立流量源枝的顺序对枝进行编号(图 31). 按照以上编号，该系统的变量(压力差、流量)可分类如下

$$
\boldsymbol{V}=\begin{bmatrix}\boldsymbol{V}_C\\ \hdashline \boldsymbol{V}_R\\ \hdashline \boldsymbol{V}_I\end{bmatrix}=\begin{bmatrix}P_1\\ P_2\\ \hdashline P_3\\ P_4\\ \hdashline P_5\end{bmatrix}\qquad
\boldsymbol{I}=\begin{bmatrix}\boldsymbol{I}_C\\ \hdashline \boldsymbol{I}_R\\ \hdashline \boldsymbol{I}_I\end{bmatrix}=\begin{bmatrix}f_1\\ f_2\\ \hdashline f_3\\ f_4\\ \hdashline f_5\end{bmatrix}
$$

由此分类可见，可以将 P_1, P_2 作为该系统的状态变量. 其次，关于 P 树的基本割集矩阵是

$$
\boldsymbol{Q}_f=\overset{\text{枝}}{\overbrace{\begin{array}{cc:ccc} 1 & 2 & 3 & 4 & 5\\ \end{array}}}\\
\boldsymbol{Q}_f=\left[\begin{array}{cc:ccc}1 & 0 & 1 & 0 & -1\\ 0 & 1 & -1 & 1 & 0\end{array}\right]\triangleq[\boldsymbol{I} \mid \boldsymbol{F}]
$$

将该矩阵中的部分矩阵 $\boldsymbol{F}$ 分解成下列分块矩阵

$$
\boldsymbol{F}=\left[\begin{array}{cc:c}1 & 0 & -1\\ -1 & 1 & 0\end{array}\right]=[\boldsymbol{F}_{CR} \mid \boldsymbol{F}_{CI}] \tag{51}
$$

此外，因贮水箱、流阻的端子特性式分别为

$$
\begin{bmatrix}f_1(t)\\ f_2(t)\end{bmatrix}=\begin{bmatrix}c_1 & 0\\ 0 & c_2\end{bmatrix}\begin{bmatrix}\dot{P}_1(t)\\ \dot{P}_2(t)\end{bmatrix}
$$

$$
\begin{bmatrix}P_3(t)\\ P_4(t)\end{bmatrix}=\begin{bmatrix}r_1 & 0\\ 0 & r_2\end{bmatrix}\begin{bmatrix}f_3(t)\\ f_4(t)\end{bmatrix}
$$

则

$$
\begin{bmatrix}c_1 & 0\\ 0 & c_2\end{bmatrix}=\boldsymbol{C}_C,\qquad \begin{bmatrix}r_1 & 0\\ 0 & r_2\end{bmatrix}=\boldsymbol{R}_R \tag{52}
$$

利用(51)、(52)式，便得到中间标准形

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & c_1 & 0 \\ -1 & 1 & 0 & c_2 \end{bmatrix}\begin{bmatrix} f_3(t) \\ f_4(t) \\ \dot{P}_1(t) \\ \dot{P}_2(t) \end{bmatrix}=\begin{bmatrix} r_1^{-1} & -r_1^{-1} \\ 0 & r_2^{-1} \\ 0 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} P_1(t) \\ P_2(t) \end{bmatrix}+\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} f_5(t) \tag{53}$$

该式左边的系数矩阵是三角形矩阵，且其对角线元素不全为零，则为正则矩阵．因此，在(53)式两边左乘以该系数矩阵的逆矩阵，或对(53)式进行适当的基本行变换后，便可以使左边的系数矩阵变成单位矩阵

$$\begin{bmatrix} f_3(t) \\ f_4(t) \\ \dot{P}_1(t) \\ \dot{P}_2(t) \end{bmatrix}=\begin{bmatrix} r_1^{-1} & -r_1^{-1} \\ 0 & r_2^{-1} \\ -c_1^{-1}r_1^{-1} & c_1^{-1}r^{-1} \\ c_2^{-1}r_1^{-1} & -c_2^{-1}(r_1^{-1}+r_2^{-1}) \end{bmatrix}\begin{bmatrix} P_1(t) \\ P_2(t) \end{bmatrix}+\begin{bmatrix} 0 \\ 0 \\ c^{-1} \\ 0 \end{bmatrix} f_5(t) \tag{54}$$

将独立流量源的端子特性式

$$f_5(t)=\hat{f}(t) \tag{55}$$

代入(54)式，取出下侧两式便得到状态方程式

$$\begin{bmatrix} \dot{P}_1(t) \\ \dot{P}_2(t) \end{bmatrix}=\begin{bmatrix} -c_1^{-1}r_1^{-1} & c_1^{-1}r_1^{-1} \\ c_2^{-1}r_1^{-1} & -c_2^{-1}(r_1^{-1}+r_2^{-1}) \end{bmatrix}\begin{bmatrix} P_1(t) \\ P_2(t) \end{bmatrix}+\begin{bmatrix} c_1^{-1} \\ 0 \end{bmatrix}\hat{f}(t) \tag{56}$$

此外，因输出 y 是流经流阻 2 的流量，即 f_4，则由(54)式得出输出方程式

$$y(t)=[0 \quad r_2^{-1}]\begin{bmatrix} P_1(t) \\ P_2(t) \end{bmatrix} \tag{57}$$

(例 4)　电气-回转运动系统的例子．

本节阐明了，引入横断变量、通过变量、代数环节、一阶横断环节等概念后，各种各样的物理系统都可以统一处理．下面再来讨论一种系统，它由几个属于不同系统的环节组成．图 32 所示系统就是其中的一个典型例子．在该系统中，设输入是独立电压源的电压 $\hat{v}$，输出 y 为马达转子的角速度．图 33 上将该系统用马达的等效模型(图 23)表示．注意，在该系统的图上，有两个连结性部分图分离存在．如图 34 所示，选取各个部分图的 N 树[1]，按

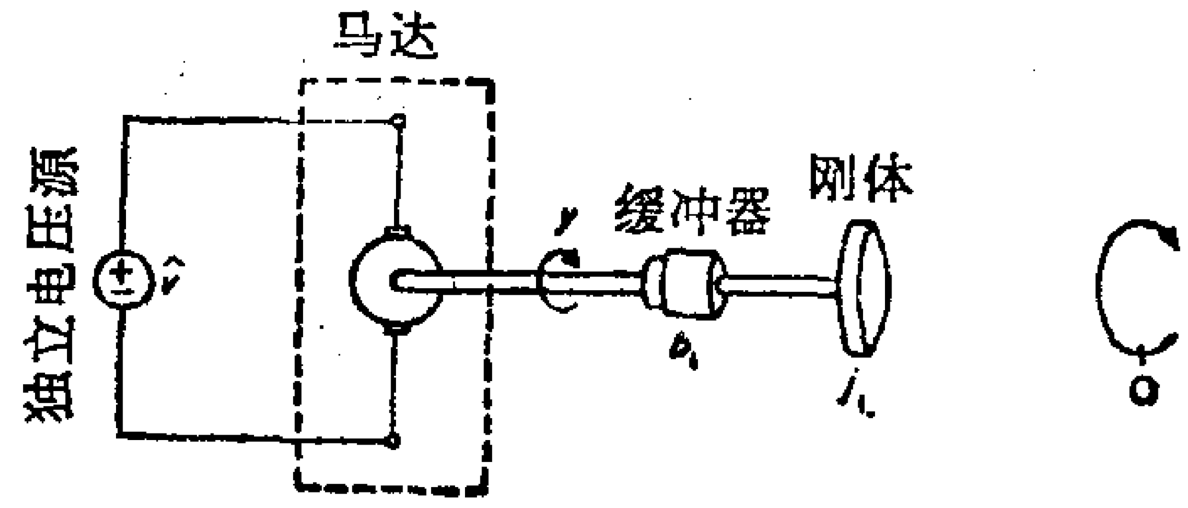


图 32　包含有马达的系统

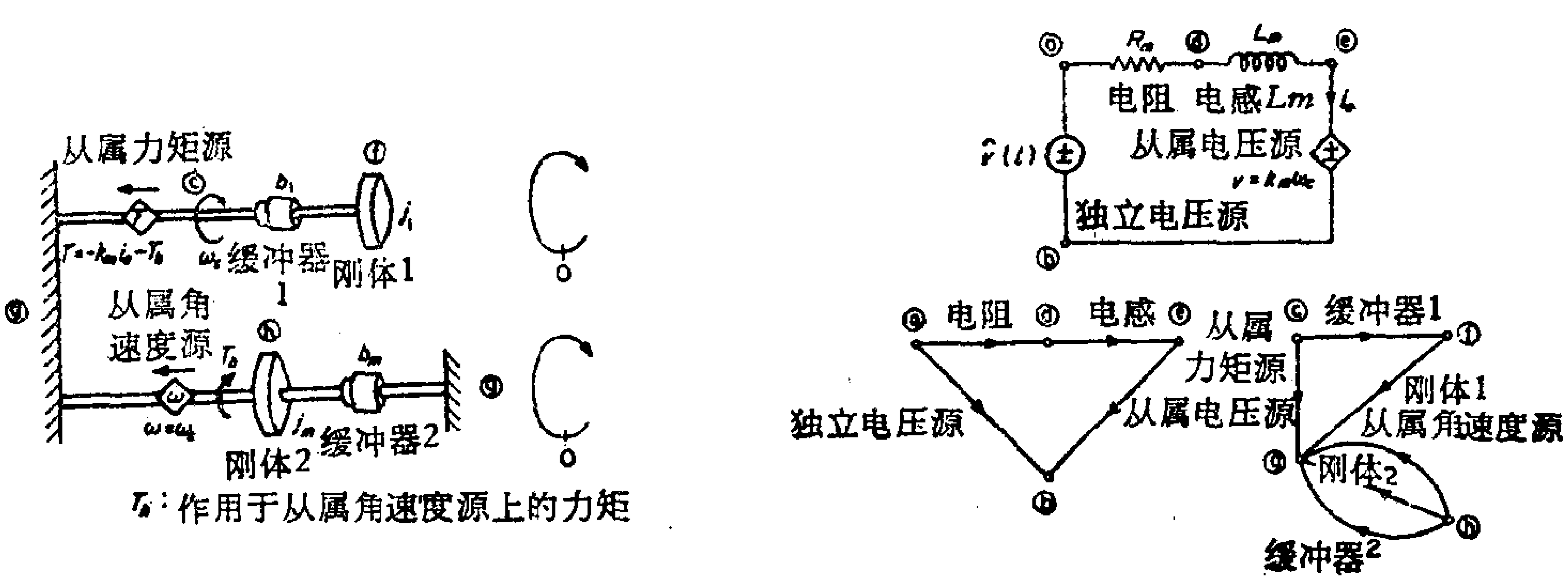


图 33　图 32 所示系统的等效模型及其图表示

1) 由几个连结性部分图构成的集合称为森(forest)．

着从属横断变量源(从属角速度源、从属电压源)枝、独立横断变量源(独立电压源)枝、包含在 N 树中的一阶横断环节(刚体 1)枝、包含在 N 树中的代数环节(缓冲器 1、电阻)枝、包含在树余中的一阶横断环节(刚体 2)枝、包含在树余中的代数环节(缓冲器 2)枝、包含在树余中的一阶通过环节(电感)枝、从属通过变量源(从属力矩源)枝的顺序对枝进行编号(图 34). 利用该编号表示各枝的变量(角速度差、力矩、电压、电流),并按 A-I-2 的符号对其进行分类

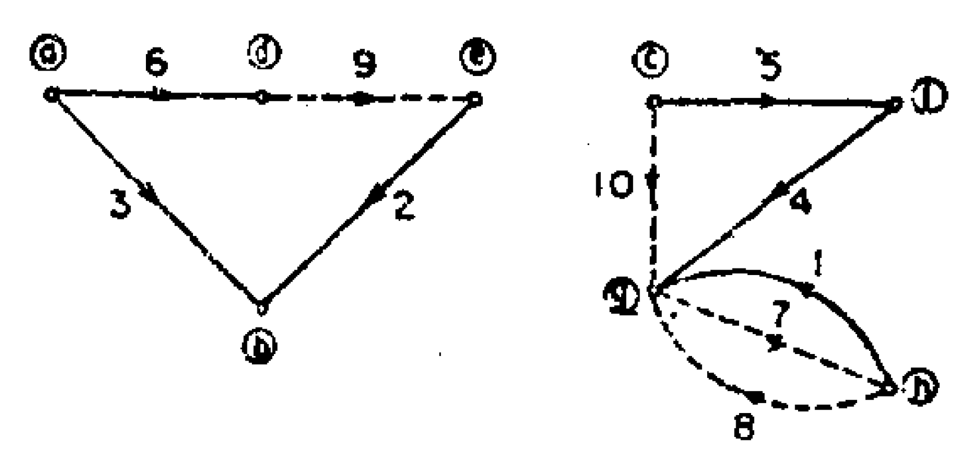


图 34 图 33 所示图的 N 树(森)及枝的编号

$$\boldsymbol{V}=\begin{bmatrix}\boldsymbol{V}_A\\ \boldsymbol{V}_V\\ \boldsymbol{V}_C\\ \boldsymbol{V}_G\\ \boldsymbol{V}_S\\ \boldsymbol{V}_R\\ \boldsymbol{V}_L\\ \boldsymbol{V}_B\end{bmatrix}=\begin{bmatrix}\omega_1\\ v_2\\ v_3\\ \omega_4\\ \omega_5\\ v_6\\ \omega_7\\ \omega_8\\ v_9\\ \omega_{10}\end{bmatrix},\qquad \boldsymbol{I}=\begin{bmatrix}\boldsymbol{I}_A\\ \boldsymbol{I}_V\\ \boldsymbol{I}_C\\ \boldsymbol{I}_G\\ \boldsymbol{I}_S\\ \boldsymbol{I}_R\\ \boldsymbol{I}_L\\ \boldsymbol{I}_B\end{bmatrix}=\begin{bmatrix}T_1\\ i_2\\ i_3\\ T_4\\ T_5\\ i_6\\ T_7\\ T_8\\ i_9\\ T_{10}\end{bmatrix}$$

由该分类可见,图 33 所示系统的补充状态变量是 ω_4、ω_7、i_9.

那么,关于 N 树的基本割集矩阵是

$$\boldsymbol{Q}_f=\begin{array}{c}\overbrace{\begin{matrix}1&2&3&4&5&6&7&8&9&10\end{matrix}}^{\text{枝}}\\ \left[\begin{array}{cccccc|cccc}1&0&0&0&0&0&1&1&0&0\\ 0&1&0&0&0&0&0&0&-1&0\\ 0&0&1&0&0&0&0&0&1&0\\ 0&0&0&1&0&0&0&0&0&1\\ 0&0&0&0&1&0&0&0&0&1\\ 0&0&0&0&0&1&0&0&-1&0\end{array}\right]\end{array}=[\boldsymbol{I}\mid\boldsymbol{F}]$$

将式中部分矩阵 $\boldsymbol{F}$ 分解成下列分块矩阵

$$\boldsymbol{F}=\left[\begin{array}{c|c|c|c}1&1&0&0\\ 0&0&-1&0\\ \hline 0&0&1&0\\ \hline 0&0&0&1\\ \hline 0&0&0&1\\ 0&0&-1&0\end{array}\right]=\left[\begin{array}{c|c|c|c}\boldsymbol{F}_{AS}&\boldsymbol{F}_{AR}&\boldsymbol{F}_{AL}&\boldsymbol{F}_{AB}\\ \hline \boldsymbol{F}_{VS}&\boldsymbol{F}_{VR}&\boldsymbol{F}_{VL}&\boldsymbol{F}_{VB}\\ \hline \boldsymbol{F}_{CS}&\boldsymbol{F}_{CR}&\boldsymbol{F}_{CL}&\boldsymbol{F}_{CB}\\ \hline \boldsymbol{0}&\boldsymbol{F}_{GR}&\boldsymbol{F}_{GL}&\boldsymbol{F}_{GB}\end{array}\right]\tag{58}$$

而且，分别得到基本割集方程式、基本回路方程式如下

$$
\begin{bmatrix} T_1(t) \\ i_2(t) \\ \hdashline i_3(t) \\ \hdashline T_4(t) \\ \hdashline T_5(t) \\ \hdashline i_6(t) \end{bmatrix}
=
\begin{bmatrix} -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}
\begin{bmatrix} T_7(t) \\ \hdashline T_8(t) \\ \hdashline i_9(t) \\ \hdashline T_{10}(t) \end{bmatrix}
\tag{59}
$$

$$
\begin{bmatrix} \omega_7(t) \\ \hdashline \omega_8(t) \\ \hdashline v_9(t) \\ \hdashline \omega_{10}(t) \end{bmatrix}
=
\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}
\begin{bmatrix} \omega_1(t) \\ v_2(t) \\ \hdashline v_3(t) \\ \hdashline \omega_4(t) \\ \hdashline \omega_5(t) \\ \hdashline v_6(t) \end{bmatrix}
\tag{60}
$$

而从属角速度源、从属电压源、从属力矩源的端子特性式分别可用

$$
\left.\begin{aligned}
\omega_1(t) &= \omega_{10}(t) \\
v_2(t) &= K_m \omega_{10}(t) \\
T_{10}(t) &= -K_m i_2(t) - T_1(t)
\end{aligned}\right\}
\tag{61}
$$

给出．将基本割集方程式(59)、基本回路方程式(60)代入上列各式的右边，将从属源的特性式仅用(59)、(60)式右边出现的变量表示，便推导出与 A-I-2 中(48)式相当的关系式

$$
\begin{bmatrix} \omega_1(t) \\ v_2(t) \\ \hdashline T_{10}(t) \end{bmatrix}
=
\begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & K_m & 0 & 0 & 0 & K_m & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & -K_m & 0 \end{bmatrix}
\begin{bmatrix} \omega_1(t) \\ v_2(t) \\ \hdashline T_{10}(t) \\ \hdashline \omega_5(t) \\ v_6(t) \\ \hdashline T_8(t) \\ \hdashline T_7(t) \\ \hdashline \omega_4(t) \\ \hdashline i_9(t) \\ \hdashline v_3(t) \end{bmatrix}
\tag{62}
$$

将该式右边系数矩阵分解成如下分块矩阵

$$
\left[\begin{array}{cc:c:cc:c:c:c:c:c} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & K_m & 0 & 0 & 0 & K_m & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & -K_m & 0 \end{array}\right]
=
\left[\begin{array}{c:c:c:c:c:c:c:c} \boldsymbol{P}_{11} & \boldsymbol{P}_{12} & \boldsymbol{P}_{13} & \boldsymbol{P}_{14} & \boldsymbol{P}_{15} & \boldsymbol{P}_{17} & \boldsymbol{P}_{18} & \boldsymbol{P}_{19} \\ \hdashline \boldsymbol{P}_{21} & \boldsymbol{P}_{22} & \boldsymbol{P}_{23} & \boldsymbol{P}_{24} & \boldsymbol{P}_{25} & \boldsymbol{P}_{27} & \boldsymbol{P}_{28} & \boldsymbol{P}_{29} \end{array}\right]
\tag{63}
$$

而且，代数环节(缓冲器、电阻)、一阶横断环节(刚体)、一阶通过环节(电感)的端子特性可分别用

$$
\begin{bmatrix} T_5(t) \\ i_6(t) \\ \hdashline \omega_8(t) \end{bmatrix}
=
\begin{bmatrix} b_1 & 0 & 0 \\ 0 & R_m^{-1} & 0 \\ 0 & 0 & b_m^{-1} \end{bmatrix}
\begin{bmatrix} \omega_5(t) \\ v_6(t) \\ \hdashline T_8(t) \end{bmatrix}
\tag{64}
$$

$$\begin{bmatrix} T_4(t) \\ \hline T_7(t) \end{bmatrix} = \begin{bmatrix} j_1 & 0 \\ 0 & j_m \end{bmatrix} \begin{bmatrix} \dot{\omega}_4(t) \\ \dot{\omega}_7(t) \end{bmatrix} \tag{65}$$

$$v_9(t) = L_m \dot{i}_9(t) \tag{66}$$

给出．把上列诸式右边的系数矩阵也分解成分块矩阵

$$\left[\begin{array}{cc|c} b_1 & 0 & 0 \\ 0 & R_m^{-1} & 0 \\ \hline 0 & 0 & b_m^{-1} \end{array}\right] = \left[\begin{array}{c|c} \boldsymbol{G}_G & \boldsymbol{0} \\ \hline \boldsymbol{0} & \boldsymbol{R}_R \end{array}\right] \tag{67}$$

$$\left[\begin{array}{c|c} j_1 & 0 \\ \hline 0 & j_m \end{array}\right] = \left[\begin{array}{c|c} \boldsymbol{C}_C & \boldsymbol{0} \\ \hline \boldsymbol{0} & \boldsymbol{C}_S \end{array}\right] \tag{68}$$

$$\boldsymbol{L}_m = \boldsymbol{L}_{LL} \tag{69}$$

利用上面得到的分块矩阵，即(58)、(63)、(67)～(69)式，便可得到中间标准形(A-I-2 中的(57)式)

$$\begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -K_m & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & -j_m & 0 & 0 \\ 0 & 0 & b_1^{-1} & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ -b_m & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & j_1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & L_m \end{bmatrix} \begin{bmatrix} \omega_1(t) \\ v_2(t) \\ T_{10}(t) \\ \omega_5(t) \\ v_6(t) \\ T_8(t) \\ \dot{\omega}_7(t) \\ \dot{\omega}_4(t) \\ \dot{i}_9(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & K_m & 0 \\ 0 & 0 & -K_m \\ 0 & 0 & 0 \\ 0 & 0 & R_m \\ 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \omega_7(t) \\ \omega_4(t) \\ i_9(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} v_3(t) \tag{70}$$

对该中间标准形进行适当的基本行变换，使其左边系数矩阵变成单位矩阵，得

$$\begin{bmatrix} \omega_1(t) \\ v_2(t) \\ T_{10}(t) \\ \omega_5(t) \\ v_6(t) \\ T_8(t) \\ \dot{\omega}_7(t) \\ \dot{\omega}_4(t) \\ \dot{i}_9(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ K_m & 0 & 0 \\ -b_1 & b_1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & R_m \\ b_m & 0 & 0 \\ -(b_1+b_m)j_m^{-1} & b_1 j_m^{-1} & K_m j_m^{-1} \\ b_1 j_1^{-1} & -b_1 j_1^{-1} & 0 \\ -K_m L_m^{-1} & 0 & -L_m^{-1} R_m \end{bmatrix} \begin{bmatrix} \omega_7(t) \\ \omega_4(t) \\ i_9(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ L_m^{-1} \end{bmatrix} v_3(t) \tag{71}$$

将独立电压源的端子特性式

$$v_3(t)=\hat{v}(t) \tag{72}$$

代入(71)式，则由下侧3行得到状态方程式

$$\begin{bmatrix}\dot{\omega}_7(t)\\ \dot{\omega}_4(t)\\ \dot{i}_9(t)\end{bmatrix}=\begin{bmatrix}-(b_1+b_m)j_m^{-1} & b_1 j_m^{-1} & K_m j_m^{-1}\\ b_1 j_1^{-1} & -b_1 j_1^{-1} & 0\\ -K_m L_m^{-1} & 0 & -L_m^{-1}R_m\end{bmatrix}\begin{bmatrix}\omega_7(t)\\ \omega_4(t)\\ i_9(t)\end{bmatrix}+\begin{bmatrix}0\\ 0\\ L_m^{-1}\end{bmatrix}\hat{v}(t) \tag{73}$$

此外，因输出是从属力矩源的角速度差 ω_{10}，则利用(60)式第4式可表示成

$$y(t)=\omega_4(t)+\omega_5(t)$$

而由(71)式的第4式可见，因 $\omega_5=\omega_7-\omega_4$，则最后得到

$$y(t)=[1\quad 0\quad 0]\begin{bmatrix}\omega_7(t)\\ \omega_4(t)\\ i_9(t)\end{bmatrix} \tag{74}$$

A-I-4 大系统和小系统

在上一章，我们把电气系统、一般物理系统视为一个动力学系统，讨论了其状态方程式的推导方法等等．从本章起，我们将对如上得到的状态方程式求解，研究系统的构造(稳定性、可控性、可观测性等)．在讨论这些内容之前，再稍用点篇幅谈一下下面几个问题．

一个动力学系统，在其相当大的情况下，能否用几个较小的动力学系统的组合表示？若能将原来的大系统有规律地分解成几个较小的系统，则可分别对分解后较小的系统进行研究，这要比原封不动地研究整个大系统方便得多．

反之，当给出几个小系统，由它们组合成一个大系统时，能否由各小系统的特性确定大系统的特性呢？例如，当各小系统的特性分别用状态方程式给出，以这些小系统的状态向量为元素作出大的状态向量时，能否立即断定它就是大系统的状态向量呢？

本章前半部分将讨论其中的第一个问题．在这里将证明，若使用表示系统环节组合状态的特殊矩阵，则能透彻地讨论系统的分解．在后半部分将讨论第二个问题．B-I-1中的分块矩阵的运算，在大系统状态方程式的推导方法中将起着重要的作用．

系统的图及布尔矩阵

在集合$\{0, 1\}$上规定如下运算规则

$$\begin{array}{lll} 0+0=0 & 0+1=1+0=1 & 1+1=1 \\ 0\times 0=0 & 0\times 1=1\times 0=0 & 1\times 1=1 \end{array} \tag{1}$$

这样定义运算规则的集合，便是一种称为布尔代数的集合．关于布尔代数的详细讨论这里从略．但是大家知道，在逻辑代数及开关电路分析等方面它已获得广泛应用．

在布尔代数上定义的矩阵，叫做布尔矩阵(Boolean matrix)[19]．因为布尔代数不是体(试与伽罗瓦体GF(2)比较)，则当对布尔矩阵使用上述各种性质时必须加以注意．当矩阵的加法、乘法按P-O-1中(3)、(4)式定义，即对于$n\times n$阶布尔矩阵$\boldsymbol{A}=[a_{ij}]$，$\boldsymbol{B}=[b_{ij}]$，令

$$\boldsymbol{AB}\triangleq\left[\sum_{l=1}^{n}a_{il}\times b_{lj}\right]$$

$$\boldsymbol{A}+\boldsymbol{B}\triangleq[a_{ij}+b_{ij}]$$

时，容易证明下列诸性质：

(i) 满足加法的交换律、结合律；

(ii) 满足乘法的结合律；

(iii) 满足加法、乘法之间的分配律．

单位矩阵$\boldsymbol{I}$及顺列矩阵$\boldsymbol{P}$(参看P-I-1)都是以1，0为元素的矩阵．将它们看成布尔矩阵，使用(1)式所示运算规则时(和实数运算的情况一样)，下列诸性质也成立：

(iv) $\boldsymbol{IA}=\boldsymbol{AI}=\boldsymbol{A}$；

(v) 左(右)乘以 $\boldsymbol{P}$，表示行(列)的交换；

(vi) $\boldsymbol{P}^T\boldsymbol{P}=\boldsymbol{P}\boldsymbol{P}^T=\boldsymbol{I}$

当 $\boldsymbol{A}$、$\boldsymbol{B}$ 均为 $m\times n$ 布尔矩阵时，其逻辑积 $\boldsymbol{A}\cap\boldsymbol{B}$ 重新定义如下

$$\boldsymbol{A}\cap\boldsymbol{B}\triangleq[a_{ij}\times b_{ij}]$$

注意：$\boldsymbol{A}\cap\boldsymbol{B}$ 和 $\boldsymbol{A}\cdot\boldsymbol{B}$ 不同.

[问题 1] 交换律、结合律对于逻辑积是否成立？

所有的元素均为 1 的 $m\times n$ 布尔矩阵用 $\boldsymbol{U}_{m\times n}$ 表示，则

$$\boldsymbol{U}_{m\times n}\cap\boldsymbol{A}=\boldsymbol{A}\cap\boldsymbol{U}_{m\times n}=\boldsymbol{A}$$

即 $\boldsymbol{U}_{m\times n}$ 是对于逻辑积的单位元. 现在，假设给出系统的信号流图. 大家知道，信号流图是一种使变量与节点相对应，用枝表示变量之间影响方式的有向图. 这里要注意，枝的方向与信号的方向一致，反方向没有信号流动. 而且，我们用 $n\times n$ 阶布尔矩阵表示 n 个节点之间的连接关系. 即认为，若有方向为从第 j 个节点到第 i 个节点的枝存在时，元素 (i,j) 为 1；而当没有枝存在，或即使有枝存在，但方向是从 i 到 j，则元素 (i,j) 为 0. 例如，在 A-I-2 开始所举的例子中，其连接关系如 A-I-2 中图 4 所示，其相应的布尔矩阵为

$$\boldsymbol{A}=\begin{bmatrix}0&0&0&0\\1&0&0&0\\0&1&0&0\\1&1&1&0\end{bmatrix},$$

这种矩阵称为所给图的节点关联矩阵或相伴矩阵[1)](association matrix). 注意，相伴矩阵总是方阵，而且很容易证明它具有下列性质(a)～(c)[50, 51]：

(a) 相伴矩阵与图一一对应. 若给出图，则其相伴矩阵唯一确定. 其逆亦真[2)]；

(b) 在相伴矩阵中，若某一行，例如第 i 行是 $\boldsymbol{0}$，则第 i 个节点是系统的输入；若第 j 列为 $\boldsymbol{0}$，则第 j 个节点是系统的输出.

(c) 设某图的相伴矩阵为 $\boldsymbol{A}$，则 $\boldsymbol{A}^T$ 与所给图中将所有的枝反向后得到的图相对应.

从第 j 个节点出发，若途中经过 $k(k\geqslant1)$ 个枝，沿该枝的方向到第 i 个节点时，称为从第 j 个节点到第 i 个节点有长度为 k 的通路存在. 特别是，当 $i=j$ 时，即后来又回到原出发节点的通路，称为回路. 此外，将长度为 0 的通路定义为停留在处于不动的节点.

(d) 设 $\boldsymbol{A}$ 为相伴矩阵，若从 j 到 i 有长度为 k 的通路存在，则 $\boldsymbol{A}^k$ 的元素 (i,j) 为 1，否则为 0. 特别是，$\boldsymbol{A}^k$ 的对角线元素表示是否有回路存在. $\boldsymbol{A}^0=\boldsymbol{I}$.

(e) $\sum\limits_{k=0}^{P}\boldsymbol{A}^k$ 表示是否有长度为 P 以下的通路存在，而且

$$\sum_{k=0}^{P}\boldsymbol{A}^k=(\boldsymbol{A}+\boldsymbol{I})^P \tag{2}$$

(f) 若图中没有回路存在，则有这样一个 $\nu(\nu\leqslant n)$ 存在，使得

$$\boldsymbol{A}^k=\boldsymbol{0}\quad k\geqslant\nu \tag{3}$$

1) 也有的书把这里定义的矩阵的转置矩阵叫做相伴矩阵，但本质上没有区别.

2) 这里假定从节点 j 到节点 i 的枝最多只有一个，若有两个以上，要预先整理成 1 个.

成立．其中 n 是图中节点的总数．

(g) 在有回路存在的图上，虽然一般没有满足(2)式的 ν 存在，但对于足够大的 N，

$$\sum_{k=0}^{N} \boldsymbol{A}^k = (\boldsymbol{A}+\boldsymbol{I})^N$$

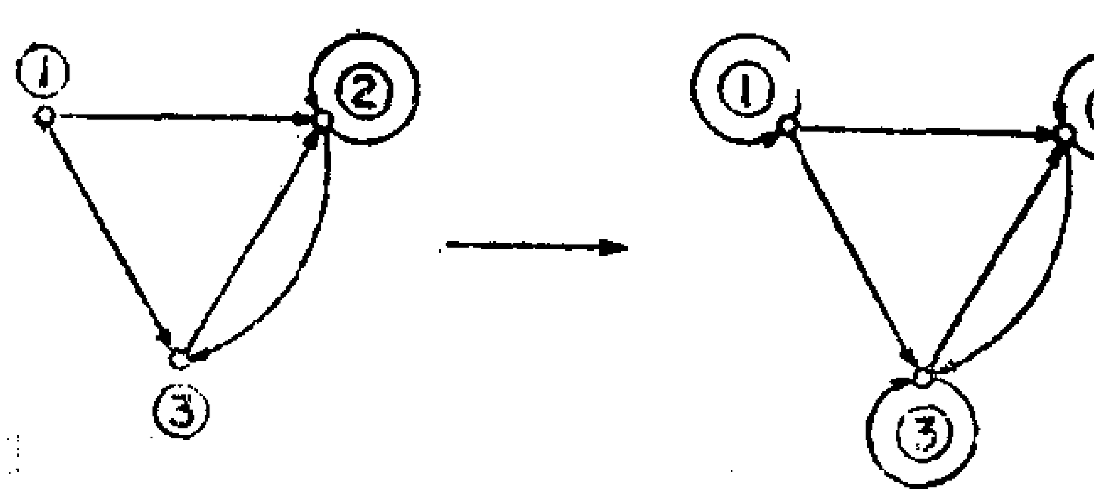
收敛于一个确定的矩阵，设该矩阵为 $\boldsymbol{R}$.

若矩阵 $\boldsymbol{R}$ 的 (i, j) 元素为 1，则从节点 j 可沿几个指定枝的方向到达节点 i，若 (i, j) 元素为 0，则从节点 j 不能到达节点 i．在这个意义上，称 $\boldsymbol{R}$ 为能达矩阵(reachability matrix)．

上面(a)～(c)是很显然的，对于(d)只说明一下 $\boldsymbol{A}^2$．$\boldsymbol{A}^2$ 的 (i, j) 元素是

$$(\boldsymbol{A}^2)_{i,j} = \sum_{l=1}^{n} a_{il} a_{lj}. \tag{4}$$

根据(1)式运算规则，当且仅当 $a_{il}a_{lj}=1$ 时，换言之，当且仅当从节点 j 到节点 l 及从节点 l 到节点 i 都有枝存在时，乘积 $a_{il}a_{lj}$ 才为 1．而且同样根据(1)式，(4)式右边至少对于一个 l，当且仅当 $a_{il}a_{lj}=1$ 时才为 1．因此，至少对于一个 l，当且仅当从 j 到 l，从 l 到 i 都有枝存在时，$(\boldsymbol{A}^2)_{i,j}$ 才为 1．亦即，若从 j 到 i 有长度为 2 的通路存在，则 $(\boldsymbol{A}^2)_{i,j}$ 为 1，否则为 0[1]．对于一般的 $\boldsymbol{A}^k (k \geqslant 2)$，同样可以采用关于 k 的归纳法加以证明．



$\boldsymbol{A}$ 的图　　$(\boldsymbol{A}+\boldsymbol{I})$的图

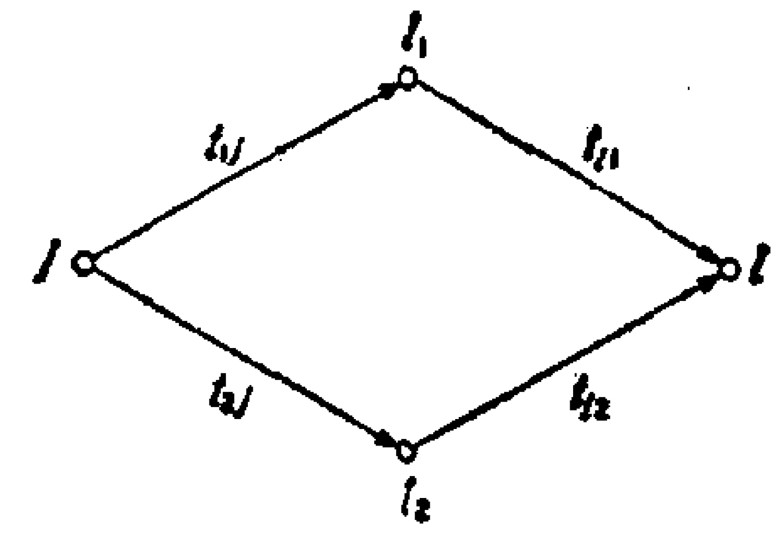
图 1

根据(d)，以及 $\sum \boldsymbol{A}^k$ 的 (i, j) 元素至少对于一个 k，当且仅当 $(\boldsymbol{A}^k)_{i,j}=1$ 时才为 1，(e)的前半部分是很显然的，下面我们来证明其后半部分，即证明 (2) 式成立．在用 $\boldsymbol{A}$ 表示的图中，将没有回路枝的节点(即 $a_{ii}=0$) 加上回路枝，而其它不变，便得到用相伴矩阵 $(\boldsymbol{A}+\boldsymbol{I})$ 表示的图(图 1)．根据(d)，在 $(\boldsymbol{A}+\boldsymbol{I})$ 的图上，若从 j 到 i 有长度为 P 的通路存在，则(2)式右边的 (i, j) 元素为 1，否则为 0．

设在 $(\boldsymbol{A}+\boldsymbol{I})$ 的图上从 j 到 i 有长度为 P 的通路存在，该通路含有 l 个"添加上去的枝"．除了这 l 个枝以外，其余仍然是从 j 到 i 的通路，长度为 $P-l$．因此，在 $\boldsymbol{A}$ 的图上，从 j 到 i 有 $k=P-l$ 个枝组成的通路．反之，若在 $\boldsymbol{A}$ 的图上从 j 到 i 有由 k 个枝 $(k \leqslant P)$ 组成的通路，则在 $(\boldsymbol{A}+\boldsymbol{I})$ 的图上有长度为 P 的通路存在(经过从 j 到 i 的 k 个枝以后，沿从 i 到 i 的枝绕 $(P-k)$ 次即可)．因此，(2)式左右两边的 (i, j) 元素相等，即表示(2)式成立．

1) 在(4)式中，设乘积 $a_{il}a_{lj}$ 对于两个 l (例如 l_1 和 l_2) 均为 1．如下面的信号流图所示，它意味着节点 j 与 i 之间有两条通路存在．在这里，系数之间偶尔有下列关系存在：

$$t_{i1}t_{1j} + t_{i2}t_{2j} = 0.$$



实际上，信号不能从 j 到达 i．但是必须注意，即使在这种情况下，用布尔矩阵分析时也会得出"从 j 到 i 有通路存在"的结论．关于下面的(e)也是同样的．

若无回路存在，则任何通路也不能两次以上通过同一节点，因此无长度为 n 的通路存在，即(f)成立.

设从 j 到 i 有长度 $k\geqslant n$ 的通路存在，因 $k\geqslant n$，则该通路两次以上通过同一节点(例如 l_1). 除掉从最先通过 l_1 到最后通过 l_1 的通路，仍然可以得到从 j 到 i 的通路. 反复进行该操作，到没有两次以上通过同一节点的通路为止，便得到长度在$(n-1)$以下的从 j 到 i 的通路. 由此可见，一般若有从 j 到 i 的通路存在，则必有一个长度为$(n-1)$的通路存在. 由此及(d)，得

$$\sum_{k=0}^{n-1} \boldsymbol{A}^k=\sum_{k=0}^{N}\boldsymbol{A}^k, \quad N\geqslant n$$

(g)得以证明.

能分解的系统

试看用相伴矩阵

$$\boldsymbol{A}_1=\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

表示的系统，其图如图 2 所示. 可见该系统实际上分离为两个子系统. 因此，该系统的分析可以在每个子系统上进行. 试将节点编号更换如下：

$$\begin{matrix} 2 & 3 & 1 & 4 \\ \downarrow & \downarrow & \downarrow & \downarrow \\ 1 & 2 & 3 & 4 \end{matrix}$$

则相伴矩阵变成下列分块对角矩阵

$$\begin{array}{l} \tilde{\boldsymbol{A}}_1=\left[\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{l} \left.\vphantom{\begin{array}{c}0\\0\end{array}}\right\} \begin{array}{l}\text{子系统 I}\\ \text{(节点 ② 和 ③)}\end{array} \\ \left.\vphantom{\begin{array}{c}0\\0\end{array}}\right\} \begin{array}{l}\text{子系统 II}\\ \text{(节点 ① 和 ④)}\end{array} \end{array} \\ \qquad\quad \text{子系统 I 子系统 II} \end{array}$$

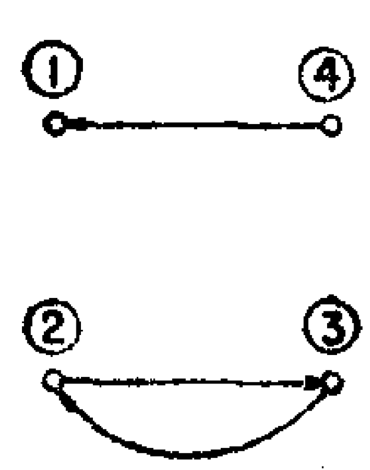


图 2

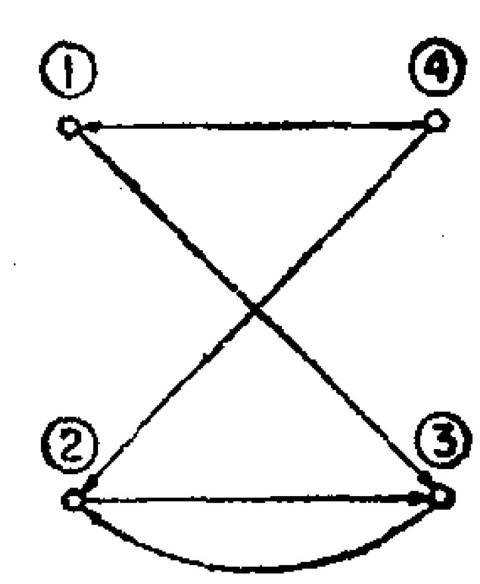


图 3

显然，该系统可以分解成两个相互无关的子系统 I 和 II.

其次，再来看看相伴矩阵

$$\boldsymbol{A}_2=\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

该矩阵表示的系统如图 3 所示．若将图中的节点分成分别由 ①、④ 和 ②、③ 组成的两个子系统，则有方向为从前者指向后者的枝，而无从后者指向前者的枝．因此，先分析由节点 ① 和 ④ 组成的子系统 II，将其结果作为由节点 ② 和 ③ 组成的子系统 I 的输入，再进行子系统 I 的分析，即将原系统看成两级结构，这样系统仍然能分解成子系统．若这里也和前面一样地更换编号，则相伴矩阵变成

$$\tilde{\boldsymbol{A}}_2=\begin{array}{c} \left[\begin{array}{cc|cc} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{l} \left.\vphantom{\begin{array}{c}0\\0\end{array}}\right\}\text{子系统 I} \\ \left.\vphantom{\begin{array}{c}0\\0\end{array}}\right\}\text{子系统 II} \end{array} \\ \underbrace{\quad\quad}_{\text{子系统 I}}\ \text{子系统 II} \end{array}$$

的分块上三角形矩阵．显然，子系统 II 对子系统 I 有影响，而相反的影响不存在．

这样，复杂的系统若能按下列中的一种方式分解：(1)能分解成几个相互无关的子系统；(2)能分解成相互之间影响有一定方向性的几个子系统、分析时往往是很方便的．由上面的例子可见，适当更换节点编号，相伴矩阵在(1)的情况下变成分块对角矩阵，在(2)的情况下变成分块三角形矩阵．因此，可以考虑下面问题

“当给定某系统时，为了适当更换节点编号，把相伴矩阵变成分块三角形矩阵(分块对角矩阵)，需要什么条件？而且更换节点编号的程序如何？”．

相伴矩阵的行及列与节点相对应．因此，将节点编号作下列更换

$$\begin{array}{cccc} i_1 & i_2 & \cdots\cdots & i_n \\ \downarrow & \downarrow & & \downarrow \\ 1 & 2 & & n \end{array} \quad ((i_1,\ i_2,\ \cdots,\ i_n)\text{是}(1,\ 2,\ \cdots,\ n)\text{的某个排列})$$

相当于在相伴矩阵上将行及列的序号作如上更换．根据性质(v)，上述更换相当于对 $\boldsymbol{A}$ 右乘以下列 $n\times n$ 阶顺列矩阵

$$\boldsymbol{P}=[\boldsymbol{e}_{i1}\boldsymbol{e}_{i2}\cdots\boldsymbol{e}_{in}],$$

左乘以 $n\times n$ 阶顺列矩阵

$$\hat{\boldsymbol{P}}=\begin{bmatrix} \boldsymbol{e}_{i1}^T \\ \boldsymbol{e}_{i2}^T \\ \vdots \\ \boldsymbol{e}_{in}^T \end{bmatrix}=\boldsymbol{P}^T$$

因此，更换节点编号使相伴矩阵作如下变化

$$\boldsymbol{A}\to\tilde{\boldsymbol{A}}\triangleq\boldsymbol{P}^T\boldsymbol{A}\boldsymbol{P}$$

则上述问题可以改述成

“当给出相伴矩阵 $\boldsymbol{A}$ 时，使 $\tilde{\boldsymbol{A}}=\boldsymbol{P}^T\boldsymbol{A}\boldsymbol{P}$ 变成分块三角形矩阵(分块对角矩阵)的顺列矩阵 $\boldsymbol{P}$ 存在的条件如何？”

若有使 $\boldsymbol{P}^T\boldsymbol{A}\boldsymbol{P}$ 变成分块对角矩阵的 $\boldsymbol{P}$ 存在，则称 $\boldsymbol{A}$ 能分离(decomposable)．而且，若有使 $\boldsymbol{P}^T\boldsymbol{A}\boldsymbol{P}$ 变成三角形矩阵的 $\boldsymbol{P}$ 存在，则称 $\boldsymbol{A}$ 是可约的(reducible)．无论用什么样的顺列矩阵都不能使 $\boldsymbol{P}^T\boldsymbol{A}\boldsymbol{P}$ 变成分块三角形矩阵时，称为 $\boldsymbol{A}$ 是既约的(irreducible)．

连结性图，强连结性图

先忽略枝的方向，考虑无向图．这时，当任意两个节点 i 和 j 之间都有连结它们的枝

存在时，该图称为是连结性的．其次，在有向图中，对于任意两个节点 i 和 j，当从 i 到 j 及从 j 到 i 均有通路存在时，称该图是强连结性的．

由相伴矩阵的性质(c)、(g)和上面的定义可见，令

$$\overline{\boldsymbol{R}} \triangleq (\boldsymbol{A}+\boldsymbol{A}^T+\boldsymbol{I})^N \quad N \geqslant n$$

时，则成为连结性图的充分和必要条件是

$$\overline{\boldsymbol{R}} = \boldsymbol{U}_{n\times n}.$$

而且也容易证明，成为强连结性图的充分和必要条件是

$$\boldsymbol{R} = \boldsymbol{U}_{n\times n} \quad \text{即} \quad \boldsymbol{R} \cap \boldsymbol{R}^T = \boldsymbol{U}_{n\times n}.$$

由图 2、图 3 的例子可见，能分离性、可约性和图的性质密切相关．

[定理 1] “图不是连结性的”和“该图的相伴矩阵 $\boldsymbol{A}$ 是能分离的”等价．

[定理 2] “图不是强连结性的”和“该图的相伴矩阵 $\boldsymbol{A}$ 是可约的”等价．

定理 1 的证明很简单，下面只证明定理 2.

（定理 2 的证明） 设 $\boldsymbol{A}$ 可约，则有使 $\widetilde{\boldsymbol{A}} = \boldsymbol{P}^T\boldsymbol{A}\boldsymbol{P}$ 成为分块三角形矩阵的 $\boldsymbol{P}$ 存在．因 $\widetilde{\boldsymbol{A}}^2$, $\widetilde{\boldsymbol{A}}^3$, …均为分块三角形矩阵，则

$$\widetilde{\boldsymbol{R}} = \sum_{k=0}^{N} \widetilde{\boldsymbol{A}}^k$$

亦为分块三角形矩阵，当然 $\boldsymbol{R}$ 具有零元素．根据性质(iv)、(vi)，因能达矩阵

$$\boldsymbol{R} = \sum_{k=0}^{N} \boldsymbol{A}_k$$

和 $\widetilde{\boldsymbol{R}}$ 之间存在下列关系

$$\widetilde{\boldsymbol{R}} = \boldsymbol{P}^T\boldsymbol{R}\boldsymbol{P}.$$

若 $\widetilde{\boldsymbol{R}}$ 具有零元素，则 $\boldsymbol{R}$ 也具有零元素，图不是强连结性的．

其逆，即“若非强连结，则可约”，将在下节根据实际分解程序予以证明．

分解程序

完全没有回路的图，求其矩阵 $\boldsymbol{P}$ 比较容易．首先，若图不具有闭回路，则必存在没有流出枝(即没有信号流出)的节点(输出节点)．其原因是，若所有节点都具有流出枝，则从某节点 j_1 出发，经由 j_1 出发的枝到节点 j_2，再经由 j_2 流出的枝到节点 j_3，…，该过程可以无限制地重复下去，但因节点数为 n，则最迟在第 n 次就回到已经通过一次的节点，即说明有回路存在．

根据相伴矩阵的性质(b)，所有元素皆为 0 的列，对应于输出节点．设这些列的编号为 i_1^1, i_2^1, …, $i_{n_1}^1$，及

$$S_1 \triangleq \{i_1^1, i_2^1, \cdots, i_{n_1}^1\}.$$

由 $\boldsymbol{A}$ 中去掉第 i_1^1, i_2^1, …, $i_{n_1}^1$ 行及第 i_1^1, i_2^1, …, $i_{n_1}^1$ 列，而余下的行及列保持原编号．例如，即使去掉第 1 行，也不能把第 2 行提到前面叫做第 1 行，仍称为原来的第 2 行．

从图上来看，去掉对应于 S_1 的行与列，相当于把输出节点及流向它们的枝都去掉．显然，这样做并不能形成回路，则余下图仍具有输出节点．因此，余下的矩阵中仍有 $\mathbf{0}$ 列存在．设这些列的编号为 i_1^2, …, $i_{n_2}^2$，及

$$S_2 \triangleq \{i_1^2, \cdots, i_{n_2}^2\}.$$

继续进行上述操作，到无剩余列为止，得

$$S_1 \cup S_2 \cup \cdots \cup S_P = \{1, 2, \cdots, n\}$$

由此可见，若按下列更换编号

$$\begin{array}{ccccc} i_1^1, & \cdots i_{n_1}^1 & i_1^2 & \cdots\cdots & i_{n_P}^P \\ \downarrow & \downarrow & \downarrow & & \downarrow \\ 1 & n_1 & n_1+1 & \cdots\cdots & n, \end{array}$$

则 $\tilde{\boldsymbol{A}}$ 变成分块三角形矩阵(实际上是三角形矩阵)．而且显然，编号的更换不是唯一的，在 $S_1, S_2, \cdots, S_P$ 的各个当中可以编成任意顺序．

现在，我们来研究一下具有回路的一般图．假定该图不是强连结性的，因而其能达矩阵 $\boldsymbol{R}$ 具有零元素．

设在矩阵 $\boldsymbol{R} \cap \boldsymbol{R}^T$ 的某一行(例如第 j_1 行)中仅第 $i_1^1, i_2^1, \cdots, i_{n_1}^1$ 个元素为 1，而余下皆为 0，及

$$S_1 \triangleq \{i_1^1, \cdots, i_{n_1}^1\}.$$

因 $\boldsymbol{R} \cap \boldsymbol{R}^T$ 的对角线元素必定为 1，则 $j_1 \in S_1$．若仅考察属于 S_1 的节点，则为强连结性的．显然，在这上面加上任何不包含在 S_1 中的其它节点后都不是强连结性的．其次，选择 $j_2 \notin S_1$ 的第 j_2 行，设该行中仅第 $i_1^2, i_2^2, \cdots, i_{n_2}^2$ 个元素为 1．若令

$$S_2 \triangleq \{i_1^2, \cdots, i_{n_2}^2\},$$

则 $j_2 \in S_2$，且 $S_1 \cap S_2 = \phi$(为何？)．以下反复进行该过程，直至下式成立为止：

$$S_1 \cup S_2 \cup \cdots \cup S_P = \{1, 2, \cdots, n\}.$$

显然

$$S_i \cap S_j = \phi \quad i \neq j.$$

若考察仅属于 $S_i (1 \leqslant i \leqslant P)$ 的节点，则为强连结性的．而且和 S_1 的情况一样，若加上不属于 S_i 的节点，则强连结性多半不成立．

若进行下列编号更换

$$\begin{array}{ccccccc} i_1^1 \cdots i_{n_1}^1 & & i_1^2 & \cdots\cdots & i_{n_2}^2 & \cdots\cdots & i_{n_P}^P \\ \downarrow \quad \downarrow & & \downarrow & & \downarrow & & \downarrow \\ 1 \cdots n_1 & & n_1+1 & \cdots & n_1+n_2 & \cdots & n \end{array}$$

则相伴矩阵化成下列分块矩阵

$$\tilde{\boldsymbol{A}} = \begin{array}{l} \left[\begin{array}{c|c|c|c} \tilde{\boldsymbol{A}}_{11} & \tilde{\boldsymbol{A}}_{12} & \cdots & \tilde{\boldsymbol{A}}_{1P} \\ \hline \tilde{\boldsymbol{A}}_{21} & \tilde{\boldsymbol{A}}_{22} & \cdots & \tilde{\boldsymbol{A}}_{2P} \\ \hline \vdots & \vdots & & \vdots \\ \hline \tilde{\boldsymbol{A}}_{P1} & \tilde{\boldsymbol{A}}_{P2} & \cdots & \tilde{\boldsymbol{A}}_{PP} \end{array}\right] \begin{array}{l} \}\text{属于 } S_1 \text{ 的节点} \\ \}\text{属于 } S_2 \text{ 的节点} \\ \quad \vdots \\ \}\text{属于 } S_P \text{ 的节点} \end{array} \\ \quad\; S_1 \quad\;\; S_2 \quad \cdots \quad S_P \end{array}$$

节点群 $S_1, \cdots, S_P$ 之间的连接关系常用 $P \times P$ 阶布尔矩阵 $\boldsymbol{B}$ 表示．设 $\boldsymbol{B}$ 的 (i, j) 元素 b_{ij} 为

$$b_{ii} = 0 \qquad i = 1, \cdots, P$$

$$b_{ij} = \begin{cases} 1 & \tilde{\boldsymbol{A}}_{ij} \neq \boldsymbol{0} \\ 0 & \tilde{\boldsymbol{A}}_{ij} = \boldsymbol{0}. \end{cases} \qquad i, j = 1, \cdots, P, \; i \neq j$$

那么，用相伴矩阵 $\boldsymbol{B}$ 表示的图没有回路存在．其原因是，例如若具有图 4(a) 那样的回

路,则表示 $\tilde{A}_{12}$ 和 $\tilde{A}_{21}$ 都不是零矩阵,而且表示从 S_1 的某个节点(例如 i_1^1)到 S_2 的某个节点(例如 i_1^2)有枝存在. 反之,从 S_2 的某个节点(例如 i_2^2)到 S_1 的某个节点(例如 i_2^1)有枝存在(图 4(b)).

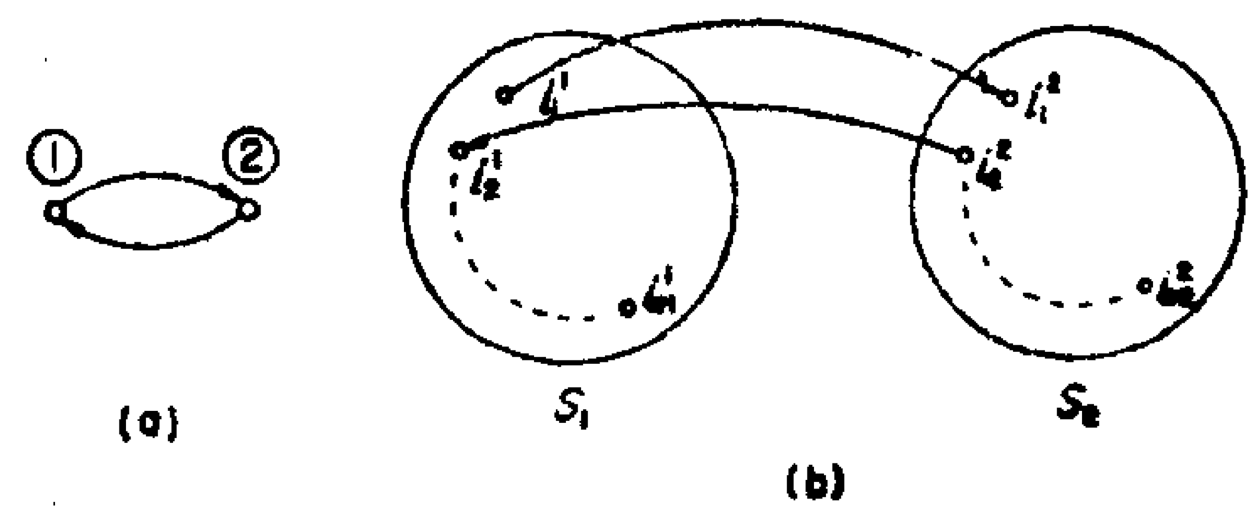


图 4

因 S_1 为强连结,则能从 S_1 的任意节点到 i_1^1. S_2 亦为强连结,从 i_1^2 也能到 S_2 的任意节点. 于是,从 S_1 的任意节点能到 S_2 的任意节点. 经过 i_2^2, i_2^1 使用同样方法证明,从 S_2 的任意节点可以到 S_1 的任意节点. 这样一来,$S_1 \cup S_2$ 成为强连结性的,与在 S_1 上加上任意节点都不是强连结性的性质相矛盾.

我们知道,$\boldsymbol{B}$ 表示没有回路的图,用前述方法更换编号后变成三角形矩阵,则可看出

$$\begin{array}{ccc} j_1 & j_2 \cdots\cdots & j_p \\ \downarrow & \downarrow & \downarrow \\ 1 & 2 \cdots\cdots & P. \end{array}$$

因此,若再改换原来图中的节点编号,使

$$\begin{array}{cccccc} i_1^{j_1} \cdots & i_{n_{j_1}}^{j_1} & i_1^{j_2} \cdots\cdots & i_{n_{j_2}}^{j_2} \cdots\cdots & i_{n_{j_p}}^{j_p} \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 1 \cdots & n_j & n_{j_1}+1 & n_{j_1}+n_{j_2} \cdots & n, \end{array}$$

则相伴矩阵 $\tilde{A}$ 变成分块三角形矩阵.

于是证明,若非强连结,则适当更换编号可以使相伴矩阵化成分块三角形矩阵.

(例题)

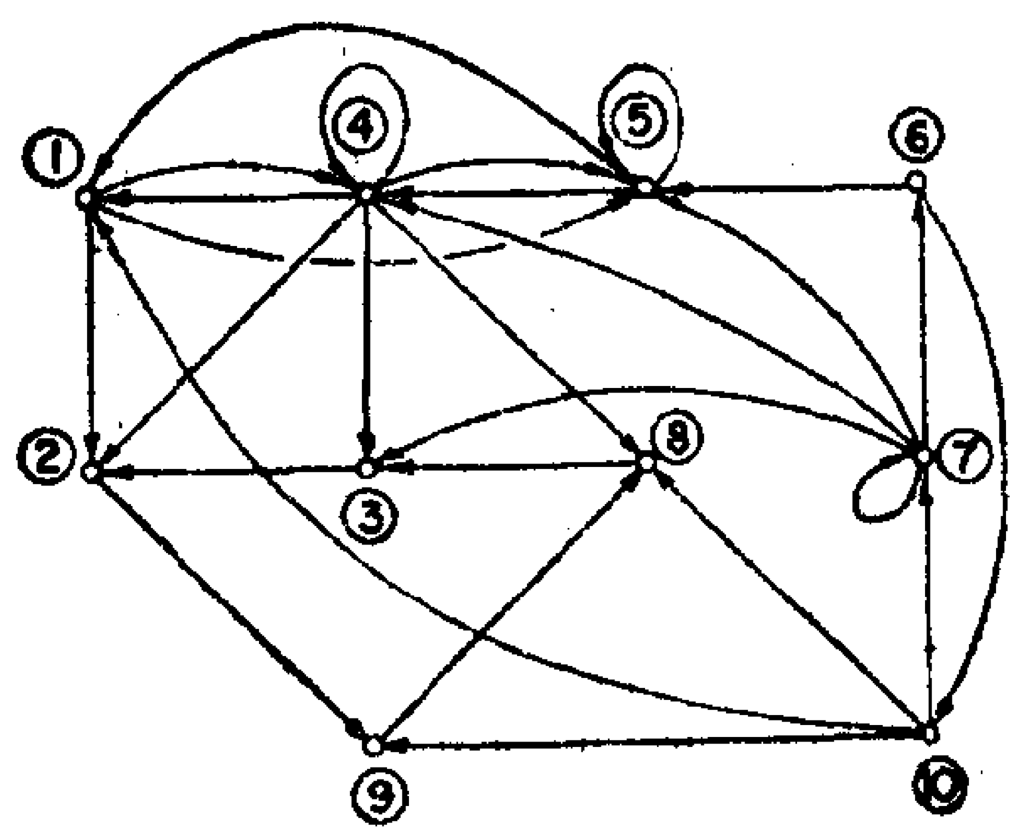


图 5

作为例子,现来看一下图 5 所示图,其相伴矩阵为 $\boldsymbol{A}$,

$$\boldsymbol{A}=\begin{bmatrix} 0&0&0&1&1&0&0&0&0&1 \\ 1&0&1&1&0&0&0&0&0&0 \\ 0&0&0&1&0&0&1&1&0&0 \\ 1&0&0&1&1&0&1&0&0&0 \\ 1&0&0&1&1&1&1&0&0&0 \\ 0&0&0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&0&1&0&0&1 \\ 0&0&0&1&0&0&0&0&1&1 \\ 0&1&0&0&0&0&0&0&0&1 \\ 0&0&0&0&0&1&0&0&0&0 \end{bmatrix}$$

$$\therefore\quad \boldsymbol{A}+\boldsymbol{I}=\begin{bmatrix}1&0&0&1&1&0&0&0&0&1\\1&1&1&1&0&0&0&0&0&0\\0&0&1&1&0&0&1&1&0&0\\1&0&0&1&1&0&1&0&0&0\\1&0&0&1&1&1&1&0&0&0\\0&0&0&0&0&1&1&0&0&0\\0&0&0&0&0&0&1&0&0&1\\0&0&0&1&0&0&0&1&1&1\\0&1&0&0&0&0&0&0&1&1\\0&0&0&0&0&1&0&0&0&1\end{bmatrix}$$

$$\boldsymbol{R}=(\boldsymbol{A}+\boldsymbol{I})^9=\begin{bmatrix}1&0&0&1&1&1&1&0&0&1\\1&1&1&1&1&1&1&1&1&1\\1&1&1&1&1&1&1&1&1&1\\1&0&0&1&1&1&1&0&0&1\\1&0&0&1&1&1&1&0&0&1\\0&0&0&0&0&1&1&0&0&1\\0&0&0&0&0&1&1&0&0&1\\1&1&1&1&1&1&1&1&1&1\\1&1&1&1&1&1&1&1&1&1\\0&0&0&0&0&1&1&0&0&1\end{bmatrix}$$

$$\boldsymbol{R}\cap\boldsymbol{R}^T=\begin{bmatrix}1&0&0&1&1&0&0&0&0&0\\0&1&1&0&0&0&0&1&1&0\\0&1&1&0&0&0&0&1&1&0\\1&0&0&1&1&0&0&0&0&0\\1&0&0&1&1&0&0&0&0&0\\0&0&0&0&0&1&1&0&0&1\\0&0&0&0&0&1&1&0&0&1\\0&1&1&0&0&0&0&1&1&0\\0&1&1&0&0&0&0&1&1&0\\0&0&0&0&0&1&1&0&0&1\end{bmatrix}$$

若选取 $i_1^1=1$，则 $S_1=\{1,\ 4,\ 5\}$，其次

若选取 $i_1^2=2$，则 $S_2=\{2,\ 3,\ 8,\ 9\}$

若选取 $i_1^3=6$，则 $S_3=\{6,\ 7,\ 10\}$

$S_1\cup S_2\cup S_3=\{1,\ \cdots,\ 10\}$，所以 $P=3$，暂且作下列置换

$$\underbrace{\begin{array}{ccc}1&4&5\\\downarrow&\downarrow&\downarrow\\1&2&3\end{array}}_{S_1}\Bigg|\underbrace{\begin{array}{cccc}2&3&8&9\\\downarrow&\downarrow&\downarrow&\downarrow\\4&5&6&7\end{array}}_{S_2}\Bigg|\underbrace{\begin{array}{ccc}6&7&10\\\downarrow&\downarrow&\downarrow\\8&9&10\end{array}}_{S_3}\Bigg|$$

于是

$$
\tilde{A}=\begin{array}{c}
\left[\begin{array}{ccc:cccc:ccc}
0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\
\hdashline
1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\
\hdashline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0
\end{array}\right]
\begin{array}{l}
\left.\begin{array}{c} \\ \\ \\ \end{array}\right\} S_1 \\
\left.\begin{array}{c} \\ \\ \\ \\ \end{array}\right\} S_2 \\
\left.\begin{array}{c} \\ \\ \\ \end{array}\right\} S_3
\end{array} \\
\begin{array}{ccc}
\underbrace{\qquad\qquad}_{S_1} & \underbrace{\qquad\qquad\qquad}_{S_2} & \underbrace{\qquad\qquad}_{S_3}
\end{array}
\end{array}
$$

矩阵 $\boldsymbol{B}$ 为

$$
\boldsymbol{B}=\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}
$$

因其第 2 列为 $\mathbf{0}$, 则将第 2 行、第 2 列消去

$$
\boldsymbol{B}=\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}
$$

第 1 列是 $\mathbf{0}$. 因此, 若作下列置换

$$
\begin{array}{ccc}
2 & 1 & 3 \\
\downarrow & \downarrow & \downarrow \\
1 & 2 & 3
\end{array},
$$

则 $\boldsymbol{B}$ 变成三角形矩阵. 实际上

$$
\boldsymbol{B}=\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \longrightarrow \tilde{\boldsymbol{B}}=\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.
$$

因此, 最后若作下列置换

$$
\begin{array}{cccc|ccc|ccc}
2 & 3 & 8 & 9 & 1 & 4 & 5 & 6 & 7 & 10 \\
\downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
\multicolumn{4}{c}{\underbrace{\qquad\qquad\qquad}_{S_2}} & \multicolumn{3}{c}{\underbrace{\qquad\qquad}_{S_1}} & \multicolumn{3}{c}{\underbrace{\qquad\qquad}_{S_3}}
\end{array}
$$

则化成分块三角形矩阵

$$
\tilde{A}=\left[\begin{array}{cccc:ccc:ccc}
0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hdashline
 & & & & 0 & 1 & 1 & 0 & 0 & 1 \\
 & \mathbf{0} & & & 1 & 1 & 1 & 0 & 1 & 0 \\
 & & & & 1 & 1 & 1 & 1 & 1 & 0 \\
\hdashline
 & & & & & & & 0 & 1 & 0 \\
 & \mathbf{0} & & & & \mathbf{0} & & 0 & 1 & 1 \\
 & & & & & & & 1 & 0 & 0
\end{array}\right]
$$

用方框图表示大系统

本章到此为止，介绍了当给出大系统时将其分解成几个小系统的方法．下面将讨论相反的问题，即对于给出的小系统(子系统)，如何由它们组合成大系统(整个系统)．在这种情况下，各子系统的特性及其结合方式如何影响整个系统特性的问题是很重要的，其中最基本的是子系统的状态方程式和整个系统状态方程式的关系．例如，是否能够说以子系统的状态向量为元素作成的大向量就是整个系统的状态向量，或者当子系统用线性标准形方程式表示时，描述整个系统的状态方程式仍然为标准形．与此密切相关的问题是，当子系统的方程式有解时，能否说整个系统的方程式也有解．以下将分别讨论这些问题．

为了表示子系统的结合方式，我们将要使用大家最熟悉的方框图．在方框图中，子系统用一个方框表示，并且假定各方框的输入输出特性已经知道．特别是，不讨论方框(子系统)内部的构造问题．本章讨论的对象仅限于非时变线性系统，其输入输出特性可以用标准形

$$\dot{x}=Ax+Bu,$$
$$y=Cx+Du$$

表示．

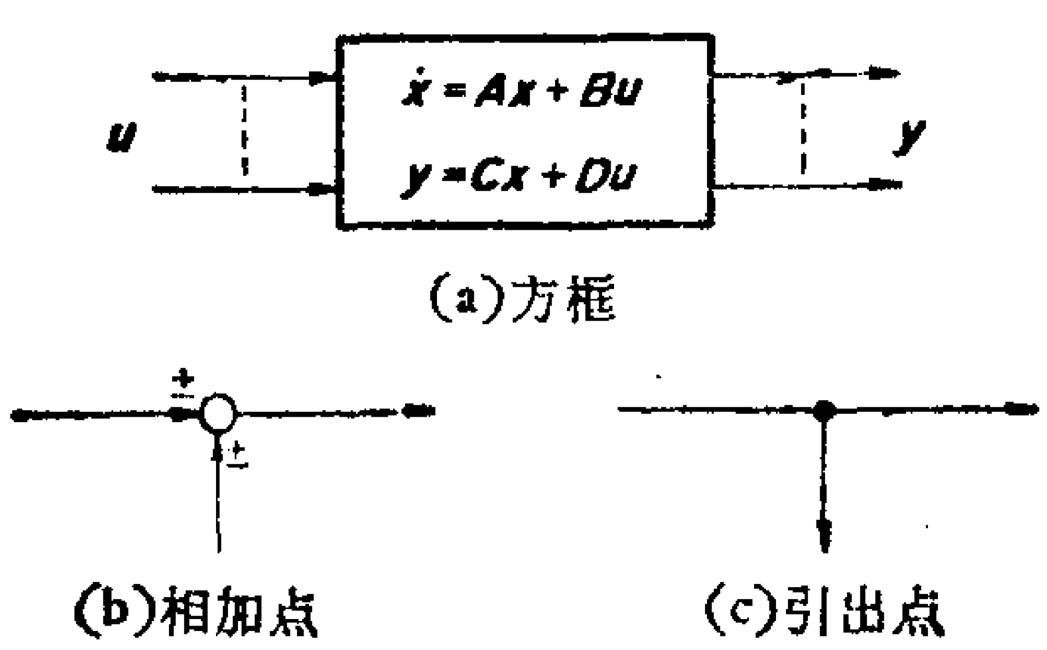


图6　方框图中的基本符号

作为方框图中的基本符号，除了方框以外，还有大家所熟悉的表示信号结合状态的相加点、引出点(参看图6)．

方框图和关联矩阵

现在，假定整个系统由 P 个方框、若干相加点，引出点组成．任意对方框进行编号，第 i 个方框用 S_i 表示．设 S_i 的输入量个数为 r_i，输出量个数为 m_i，整个系统的输入量个数为 r，输出量个数为 m．在该系统中，假定顺着从 S_j 的第 k 个输出引出的箭头，经过若干个相加点和引出点到达 S_i 的第 l 个输入 u_{il}．这时称为 y_{jk} 接到 u_{il}[1]．而且，当输入信号偶数次带着负号经过相加点时，称为正接，奇数次时称为负接．同样，在整个系统的输入和子系统的输入之间，子系统的输出和整个系统的输出之间，整个系统的输入和输出之间，也可以考虑正、负连接关系．这些连接关系可以用下述矩阵表示．

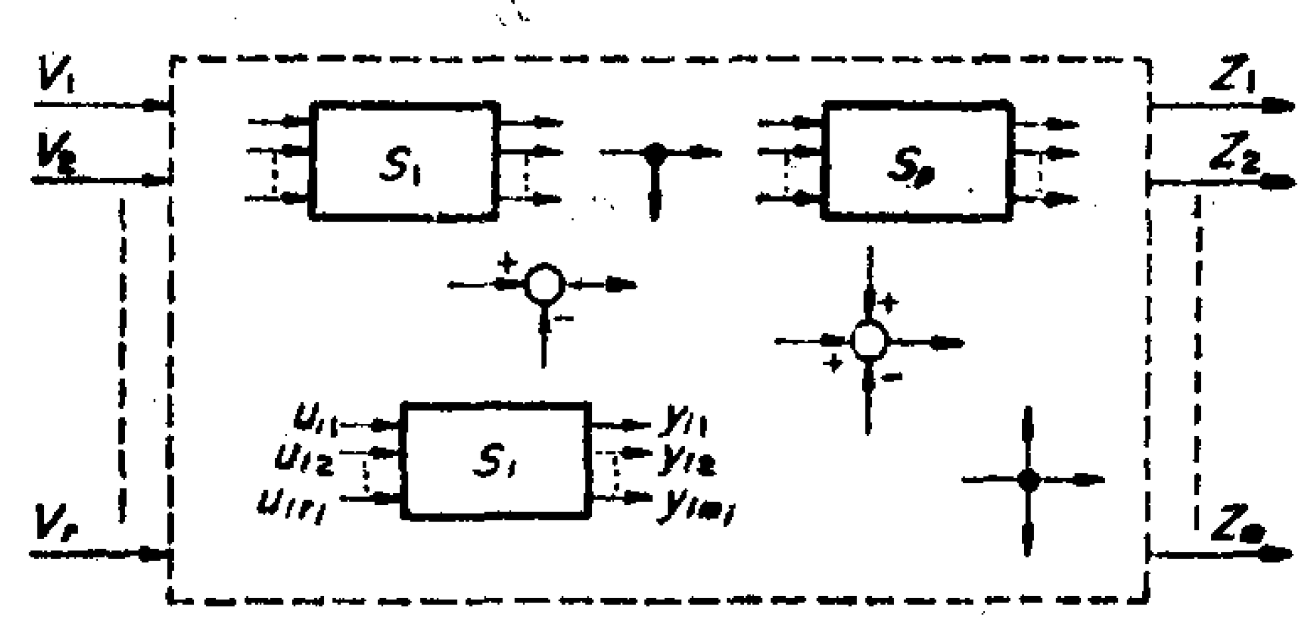


图7　方框图

表示方框之间连接关系(方框的输出和方框的输入之间的关系)的方框——方框关联矩阵 $\boldsymbol{F}$，由 P 个分块矩阵组成．

1) 这时假定从 y_{jk} 到 u_{il} 的通路只有一个．而且，在研究子系统输入和整个系统输入、子系统输出和整个系统输出、整个系统输入和输出之间的关系时，不一定都认为通路只有一个．这些假定之所以必要，是为了使连接关系的研究得到简化，它不是本质性的假定．对于多输入-单输出的方框处理相加点、单输入-多输出的方框处理引出点，这种假定就不是必要的．

$$F=\begin{bmatrix} F_{11} & F_{12} & \cdots & F_{1P} \\ F_{21} & F_{22} & \cdots & F_{2P} \\ \vdots & \vdots & & \vdots \\ F_{P1} & F_{P2} & \cdots & F_{PP} \end{bmatrix}$$

式中 ij 子块 $\boldsymbol{F}_{ij}$ 的维数是 $r_i\times m_j$，其元素决定于 S_j 的输出如何接到 S_i 的输入，即 $\boldsymbol{F}_{ij}$ 的 kl 元素 $f_{ij}(k, l)$ 按如下确定.

$$f_{ij}(k, l)=\begin{cases} 1: & \text{当 } S_j \text{ 的第 } l \text{ 个输出正接到 } S_i \text{ 的第 } k \text{ 个输入时;} \\ -1: & \text{当 } S_j \text{ 的第 } l \text{ 个输出负接到 } S_i \text{ 的第 } k \text{ 个输入时;} \\ 0: & \text{当 } S_j \text{ 的第 } l \text{ 个输出不接到 } S_i \text{ 的第 } k \text{ 个输入时.} \end{cases}$$

表示整个系统输入和方框(子系统)之间连接关系的输入-方框关联矩阵 $\boldsymbol{G}$，由 P 个子块矩阵组成

$$G=\begin{bmatrix} G_1 \\ G_2 \\ \vdots \\ G_P \end{bmatrix}$$

式中第 i 个子块 $\boldsymbol{G}_i$ 的维数为 $r_i\times r$，其 kl 元素 $g_i(k, l)$ 表示整个系统输入如何接到 S_i 的输入，按下式确定

$$g_i(k, l)=\begin{cases} 1: & \text{当整个系统的第 } l \text{ 个输入正接到 } S_i \text{ 的第 } k \text{ 个输入时;} \\ -1: & \text{当整个系统的第 } l \text{ 个输入负接到 } S_i \text{ 的第 } k \text{ 个输入时;} \\ 0: & \text{当整个系统的第 } l \text{ 个输入不接到 } S_i \text{ 的第 } k \text{ 个输入时.} \end{cases}$$

表示方框(子系统)输出和整个系统输出之间连接关系的方框-输出的关联矩阵 $\boldsymbol{J}$，也由 P 个子块矩阵组成

$$J=[J_1J_2\cdots J_P].$$

第 j 个子块矩阵的维数为 $m\times m_j$，其 k, l 元素 $j_j(k, l)$ 表示 S_j 的输出如何接到整个系统输出，按如下确定

$$j_j(k, l)=\begin{cases} 1: & \text{当 } S_j \text{ 的第 } l \text{ 个输出正接到整个系统的第 } k \text{ 个输出时;} \\ -1: & \text{当 } S_j \text{ 的第 } l \text{ 个输出负接到整个系统的第 } k \text{ 个输出时;} \\ 0: & \text{当 } S_j \text{ 的第 } l \text{ 个输出不接到整个系统的第 } k \text{ 个输出时.} \end{cases}$$

表示整个系统输入-输出之间关系的输入-输出关联矩阵 $\boldsymbol{K}$ 的维数是 $m\times r$，其 k, l 元素 $k(k, l)$ 表示整个系统输入如何接到整个系统的输出，按如下确定

$$k(k, l)=\begin{cases} 1: & \text{当整个系统第 } l \text{ 个输入正接到整个系统第 } k \text{ 个输出时;} \\ -1: & \text{当整个系统第 } l \text{ 个输入负接到整个系统第 } k \text{ 个输出时;} \\ 0: & \text{当整个系统第 } l \text{ 个输入不接到整个系统第 } k \text{ 个输出时.} \end{cases}$$

使用上面的 $\boldsymbol{F}$、$\boldsymbol{G}$、$\boldsymbol{J}$、$\boldsymbol{K}$，可以将整个系统的输入、输出，方框(子系统)的输入输出之间的关系表示如下. 这里设 $\boldsymbol{u}_i(i=1, 2, \cdots, P)$ 是表示 S_i 输入的 r_i 维实列向量，$\boldsymbol{y}_i(i=1, 2, \cdots, P)$ 是表示 S_i 输出的 m_i 维实列向量，$\boldsymbol{v}$ 是表示整个系统输入的 r 维实列向量，$\boldsymbol{z}$ 是表示整个系统输出的 m 维实列向量，则

$$\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_P \end{bmatrix}=\begin{bmatrix} F_{11} & F_{12} & \cdots & F_{1P} \\ F_{21} & F_{22} & \cdots & F_{2P} \\ \vdots & \vdots & & \vdots \\ F_{P1} & F_{P2} & \cdots & F_{PP} \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_P \end{bmatrix}+\begin{bmatrix} G_1 \\ G_2 \\ \vdots \\ G_P \end{bmatrix}v \tag{5}$$

$$
\boldsymbol{Z}=[\boldsymbol{J}_1\ \ \boldsymbol{J}_2\ \cdots\ \boldsymbol{J}_P]\begin{bmatrix}\boldsymbol{y}_1\\ \boldsymbol{y}_2\\ \vdots\\ \boldsymbol{y}_P\end{bmatrix}+K\boldsymbol{V} \tag{6}
$$

（例 1） 在图 8 所示方框图表示的系统上，试求出整个系统输入、输出，方框输入输出之间的关系.

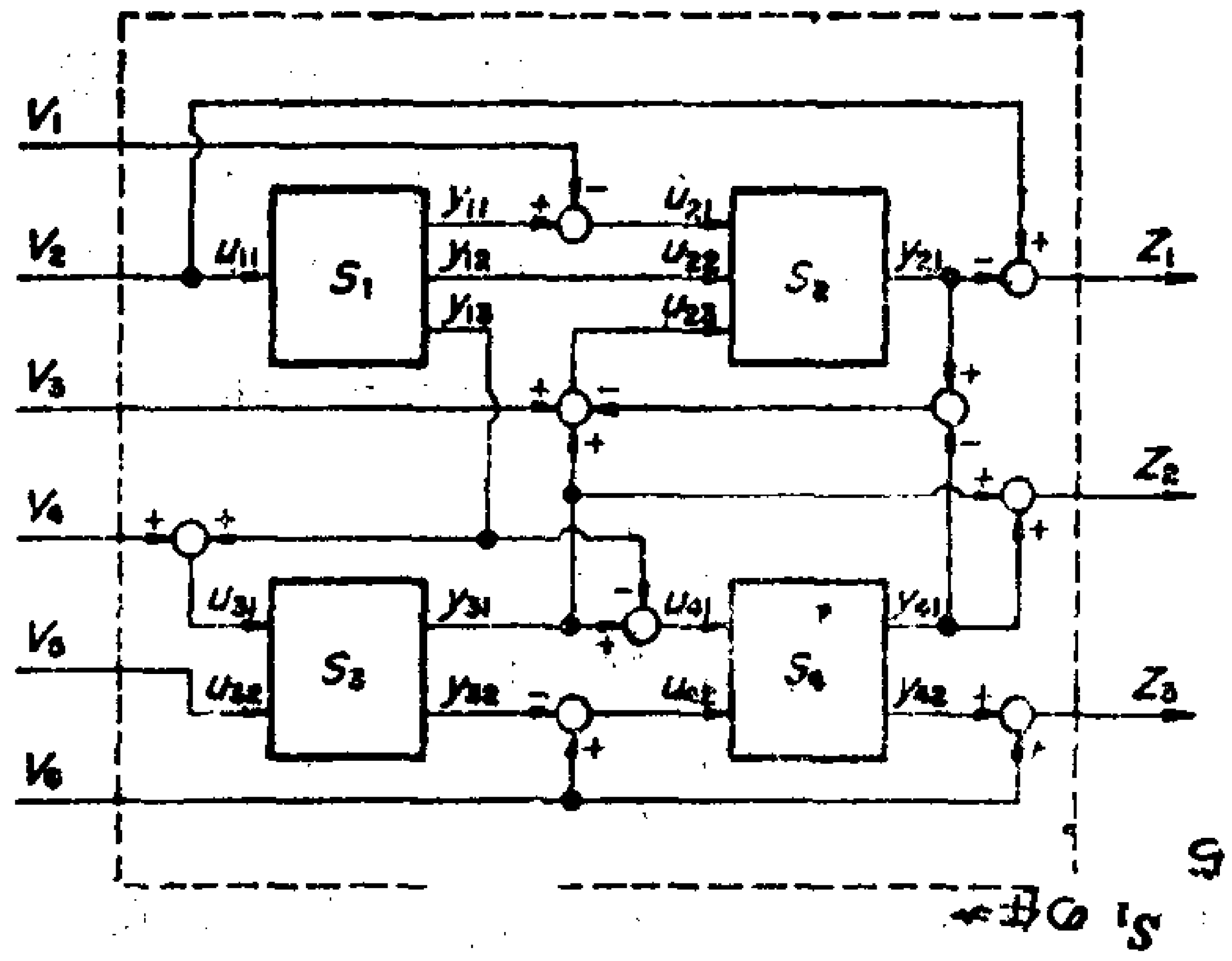


图 8　用方框图表示的例子

在该系统中，关联矩阵 $\boldsymbol{F}$、$\boldsymbol{G}$、$\boldsymbol{J}$、$\boldsymbol{K}$ 分别求出如下

$$
\boldsymbol{F}=\begin{bmatrix}\boldsymbol{F}_{11} & \boldsymbol{F}_{12} & \boldsymbol{F}_{13} & \boldsymbol{F}_{14}\\ \boldsymbol{F}_{21} & \boldsymbol{F}_{22} & \boldsymbol{F}_{23} & \boldsymbol{F}_{24}\\ \boldsymbol{F}_{31} & \boldsymbol{F}_{32} & \boldsymbol{F}_{33} & \boldsymbol{F}_{34}\\ \boldsymbol{F}_{41} & \boldsymbol{F}_{42} & \boldsymbol{F}_{43} & \boldsymbol{F}_{44}\end{bmatrix}=\left[\begin{array}{ccc|c|cc|cc}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ \hline
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & -1 & 1 & 0 & 1 & 0\\ \hline
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ \hline
0 & 0 & -1 & 0 & 1 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & -1 & 0 & 0
\end{array}\right]
$$

$$
\boldsymbol{G}=\begin{bmatrix}\boldsymbol{G}_1\\ \boldsymbol{G}_2\\ \boldsymbol{G}_3\\ \boldsymbol{G}_4\end{bmatrix}=\left[\begin{array}{cccccc}
0 & 1 & 0 & 0 & 0 & 0\\ \hline
-1 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 1 & 0 & 0 & 0\\ \hline
0 & 0 & 0 & 1 & 0 & 0\\
0 & 0 & 0 & 0 & 1 & 0\\ \hline
0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 1
\end{array}\right]
$$

$$
\boldsymbol{J}=[\boldsymbol{J}_1\ \ \boldsymbol{J}_2\ \ \boldsymbol{J}_3\ \ \boldsymbol{J}_4]=\left[\begin{array}{ccc|c|cc|cc}
0 & 0 & 0 & -1 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{array}\right]
$$

$$
\boldsymbol{K}=\begin{bmatrix}0 & 1 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 1\end{bmatrix}.
$$

因此,整个系统输入、输出,方框输入、输出之间的关系可表示如下

$$
\begin{bmatrix} u_{11} \\ \hline u_{21} \\ u_{22} \\ u_{23} \\ \hline u_{31} \\ u_{32} \\ \hline u_{41} \\ u_{42} \end{bmatrix}
=
\left[\begin{array}{ccc|c|cc|cc}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & 1 & 0 & 1 & 0 \\ \hline
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline
0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -1 & 0 & 0
\end{array}\right]
\begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ \hline y_{21} \\ \hline y_{31} \\ y_{32} \\ \hline y_{41} \\ y_{42} \end{bmatrix}
+
\left[\begin{array}{cccccc}
0 & 1 & 0 & 0 & 0 & 0 \\ \hline
-1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\ \hline
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 \\ \hline
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1
\end{array}\right]
\begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ V_5 \\ V_6 \end{bmatrix}
$$

$$
\begin{bmatrix} Z_1 \\ Z_2 \\ Z_3 \end{bmatrix}
=
\left[\begin{array}{ccc|c|cc|cc}
0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{array}\right]
\begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ \hline y_{21} \\ \hline y_{31} \\ y_{32} \\ \hline y_{41} \\ y_{42} \end{bmatrix}
+
\begin{bmatrix}
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ V_5 \\ V_6 \end{bmatrix}
$$

状态方程式的推导方法

下面我们来讨论，在用标准形状态方程式及输出方程式表示的动态方框及线性放大器方框构成的方框图上，如何选择状态变量及能否用标准形表示整个系统的问题．设线性放大器方框的输入为 $\boldsymbol{u}$，输出为 $\boldsymbol{y}$，则用适当的矩阵 $\boldsymbol{D}$ 可以表示成

$$\boldsymbol{y}=\boldsymbol{D}\boldsymbol{u}$$

它可以看成是零维标准形,即状态向量的维数是零的标准形,故以下对动态方框和放大器方框不再加以区别,认为所有方框的输入输出特性均可用下列标准形

$$\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{u},\quad \boldsymbol{y}=\boldsymbol{C}\boldsymbol{x}+\boldsymbol{D}\boldsymbol{u}$$

给出．

设方框 $S_i(i=1,\ 2,\ \cdots,\ P)$ 的输入输出特性可用标准形

$$\dot{\boldsymbol{x}}_i=\boldsymbol{A}_i\boldsymbol{x}_i+\boldsymbol{B}_i\boldsymbol{u}_i \tag{7a}$$

$$\boldsymbol{y}_i=\boldsymbol{C}_i\boldsymbol{x}_i+\boldsymbol{D}_i\boldsymbol{u}_i \tag{7b}$$

给出．式中 $\boldsymbol{x}_i$ 是 n_i 维实列向量,表示 S_i 的状态．如上所述,注意在放大器方框中 $n_i=0$.

对应于状态、输入、输出各自的维数，A_i 是 $n_i\times n_i$，B_i 是 $n_i\times r_i$，C_i 是 $m_i\times n_i$，D_i 是 $m_i\times r_i$ 的实矩阵．所有各方框的输入输出特性归纳如下

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \vdots\\ \dot{x}_P\end{bmatrix}=\begin{bmatrix}A_1 & 0\cdots 0\\ 0 & A_2\cdots 0\\ \vdots & \vdots\ \ddots\ \vdots\\ 0 & 0\cdots A_P\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_P\end{bmatrix}+\begin{bmatrix}B_1 & 0\cdots 0\\ 0 & B_2\cdots 0\\ \vdots & \vdots\ \ddots\ \vdots\\ 0 & 0\cdots B_P\end{bmatrix}\begin{bmatrix}u_1\\ u_2\\ \vdots\\ u_P\end{bmatrix}\tag{8a}$$

$$\begin{bmatrix}y_1\\ y_2\\ \vdots\\ y_P\end{bmatrix}=\begin{bmatrix}C_1 & 0\cdots 0\\ 0 & C_2\cdots 0\\ \vdots & \vdots\ \ddots\ \vdots\\ 0 & 0\cdots C_P\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_P\end{bmatrix}+\begin{bmatrix}D_1 & 0\cdots 0\\ 0 & D_2\cdots 0\\ \vdots & \vdots\ \ddots\ \vdots\\ 0 & 0\cdots D_P\end{bmatrix}\begin{bmatrix}u_1\\ u_2\\ \vdots\\ u_P\end{bmatrix}\tag{8b}$$

为了使上列方程式及整个系统输入输出、方框输入输出的关系式(5)、(6)表示简单，引入下列实向量和实矩阵

$$X\triangleq\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_P\end{bmatrix},\quad U\triangleq\begin{bmatrix}u_1\\ u_2\\ \vdots\\ u_P\end{bmatrix},\quad Y\triangleq\begin{bmatrix}y_1\\ y_2\\ \vdots\\ y_P\end{bmatrix},$$

$$A\triangleq\begin{bmatrix}A_1 & 0\cdots 0\\ 0 & A_2\cdots 0\\ \vdots & \vdots\ \ddots\ \vdots\\ 0 & 0\cdots A_P\end{bmatrix},\quad B=\begin{bmatrix}B_1 & 0\cdots 0\\ 0 & B_2\cdots 0\\ \vdots & \vdots\ \ddots\ \vdots\\ 0 & 0\cdots B_P\end{bmatrix},$$

$$C\triangleq\begin{bmatrix}C_1 & 0\cdots 0\\ 0 & C_2\cdots 0\\ \vdots & \vdots\ \ddots\ \vdots\\ 0 & 0\cdots C_P\end{bmatrix},\quad D=\begin{bmatrix}D_1 & 0\cdots 0\\ 0 & D_2\cdots 0\\ \vdots & \vdots\ \ddots\ \vdots\\ 0 & 0\cdots D_P\end{bmatrix}.$$

于是，(5)、(6)、(8a)、(8b)式可以分别写成

$$U=FY+GV\tag{5$'$}$$

$$Z=JY+KV\tag{6$'$}$$

$$\dot{X}=AX+BU\tag{8a$'$}$$

$$Y=CX+DU\tag{8b$'$}$$

因上列各式完全表示了整个系统，则能否用标准形方程式表示整个系统的问题就变成能否推导出与上列各式等价的标准形．在讨论该问题时，A-I-2 中讲过的中间标准形是很有用的．

将(5)$'$、(6)$'$、(8)$'$ 归纳后，写成中间标准形

$$\begin{bmatrix}I & -F & 0 & 0\\ -D & I & 0 & 0\\ 0 & -J & I & 0\\ -B & 0 & 0 & I\end{bmatrix}\begin{bmatrix}U\\ Y\\ Z\\ \dot{X}\end{bmatrix}=\begin{bmatrix}0\\ C\\ 0\\ A\end{bmatrix}X+\begin{bmatrix}G\\ 0\\ K\\ 0\end{bmatrix}V\tag{9}$$

由该式可见，能够以 X 为状态变量，用标准形完全表示整个系统的充分和必要条件是，上式左边系数矩阵是正则的．根据矩阵的正则性和行列式值的关系（P-I-2 中定理 1），以及分块三角形矩阵整个行列式值与其对角子块行列式值的关系(B-I-1 中定理 1)，

左边系数矩阵是正则的与

$$\det\begin{bmatrix} I & -F \\ -D & I \end{bmatrix} \neq 0$$

等价．而且，由 B-I-1 中定理 2，因

$$\det\begin{bmatrix} I & -F \\ -D & I \end{bmatrix} = \det[I-DF] = \det[I-FD],$$

故得如下定理．

［定理 3］ 以各方框状态变量的直积为状态变量，能够用标准形状态方程式及输出方程式的形式完全表示整个系统的充分和必要条件是

$$\det[I-DF]\ (=\det[I-FD]) \neq 0. \tag{10}$$

根据该定理，能否以 X 为状态变量，用标准形完全表示整个系统，仅与方框和方框之间的连接关系（即矩阵 F）及方框输入输出之间直接通路 (direct path) 的放大特性（即矩阵 D）有关．

若已知(9)式左边系数矩阵是正则的，根据 B-I-1 定理 3 求出其逆矩阵，左乘到(9)式两边，便得到和它等价的方程式

$$\begin{bmatrix} U \\ Y \\ Z \\ \dot{X} \end{bmatrix} = \begin{bmatrix} (I-FD)^{-1}FC \\ (I-DF)^{-1}C \\ J(I-DF)^{-1}C \\ A+B(I-FD)^{-1}FC \end{bmatrix} X + \begin{bmatrix} (I-FD)^{-1}G \\ (I-DF)^{-1}DG \\ K+J(I-DF)^{-1}DG \\ B(I-FD)^{-1}G \end{bmatrix} V$$

由该式下半部分可以得到关于整个系统输入输出的标准形

$$\dot{X} = \{A+B(I-FD)^{-1}FC\}X + B(I-FD)^{-1}GV \tag{11a}$$

$$Z = J(I-DF)^{-1}CX + \{K+J(I-DF)^{-1}DG\}V. \tag{11b}$$

（例 2） 现在讨论一下由两个单输入单输出方框组成的反馈系统（图 9）．由该系统的关联矩阵 F、G、J、K 求出整个系统输入输出、子系统的输入输出之间的关系如下

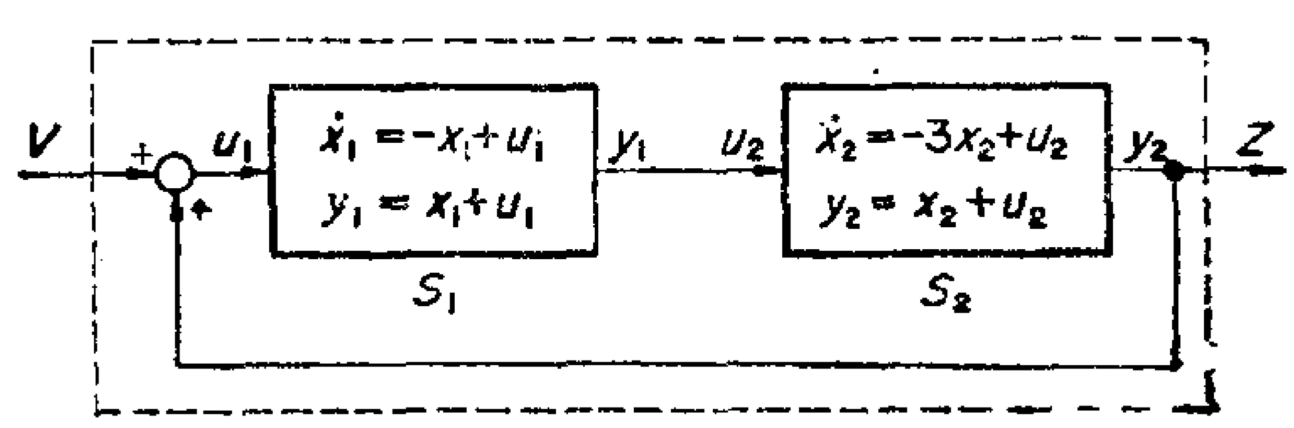


图 9 由单输入单输出方框组成的反馈系统

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \underset{F}{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \underset{G}{\begin{bmatrix} 1 \\ 0 \end{bmatrix}} V \tag{12}$$

$$Z = [0 \quad 1]\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + [0]V \tag{13}$$

另外，将方框的输入输出特性整理如下

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \underset{A}{\begin{bmatrix} -1 & 0 \\ 0 & -3 \end{bmatrix}} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \underset{B}{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \tag{14a}$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \underset{\boldsymbol{C}}{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \underset{\boldsymbol{D}}{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \tag{14b}$$

由上列各式得
$$\boldsymbol{I}-\boldsymbol{DF}=\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix},$$
$\det[\boldsymbol{I}-\boldsymbol{DF}]=0$. 因此，不能以$[x_1, x_2]^T$为状态变量用标准形完全表示整个系统.

下面，我们将一般性地讨论$\boldsymbol{I}-\boldsymbol{DF}$为正则($\boldsymbol{I}-\boldsymbol{FD}$为正则)时的连接条件. 在讨论连接关系时，前面讲过的布尔矩阵分析方法很有用. 现定义两个布尔矩阵$\boldsymbol{D}^*$、$\boldsymbol{F}^*$如下

$$\boldsymbol{D}^*=\begin{bmatrix} d_{11}^* & 0 & \cdots & 0 \\ 0 & d_{22}^* & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{pp}^* \end{bmatrix}$$

$$d_{ii}^* \triangleq \begin{cases} 1: & S_i \text{ 中有直接通路，即 } \boldsymbol{D}_i \neq \boldsymbol{0} \text{ 时;} \\ 0: & S_i \text{ 中没有直接通路，即 } \boldsymbol{D}_i = \boldsymbol{0} \text{ 时,} \end{cases}$$

$$\boldsymbol{F}^*=\begin{bmatrix} f_{11}^* & f_{12}^* & \cdots & f_{1p}^* \\ f_{21}^* & f_{22}^* & \cdots & f_{2p}^* \\ \vdots & \vdots & \ddots & \vdots \\ f_{p1}^* & f_{p2}^* & \cdots & f_{pp}^* \end{bmatrix}$$

$$f_{ij}^* \triangleq \begin{cases} 1: & S_j \text{ 的输出接到 } S_i \text{ 的输入，即 } \boldsymbol{F}_{ij} \neq \boldsymbol{0} \text{ 时;} \\ 0: & S_j \text{ 的输出不接到 } S_i \text{ 的输入，即 } \boldsymbol{F}_{ij} = \boldsymbol{0} \text{ 时.} \end{cases}$$

利用这些矩阵可以求出$\boldsymbol{I}-\boldsymbol{DF}$为正则矩阵的充分条件，由定理 3 可以得到下列系.

[系 1] 若有使
$$(\boldsymbol{D}^*\boldsymbol{F}^*)^N=\boldsymbol{0}$$
成立的正整数$N \leqslant P$存在，则可以以各方框状态变量的直积为状态变量，用标准形状态方程式和输出方程式完全地表示整个系统.

（证明） 由$\boldsymbol{D}^*$、$\boldsymbol{F}^*$的定义可见，$(\boldsymbol{D}^*\boldsymbol{F}^*)^N$的$ij$元素为 0 则表示$(\boldsymbol{DF})^N$的$ij$子块是零矩阵[1]. 因此，$(\boldsymbol{D}^*\boldsymbol{F}^*)^N=\boldsymbol{0}$意味着$(\boldsymbol{DF})^N=\boldsymbol{0}$. 现假定$\boldsymbol{I}-\boldsymbol{DF}$不是正则的. 根据 B-I-4 中定理 3，则得$(\boldsymbol{I}-\boldsymbol{DF})\boldsymbol{x}_0=\boldsymbol{0}$，即有使$\boldsymbol{x}_0=\boldsymbol{DF}\boldsymbol{x}_0$的$\left(\sum_{i=1}^{P} m_i\right)$维实列向量$\boldsymbol{x}_0(\neq\boldsymbol{0})$存在. 对该式左乘以$\boldsymbol{DF}$，得$\boldsymbol{DF}\boldsymbol{x}_0=(\boldsymbol{DF})^2\boldsymbol{x}_0$，因而$\boldsymbol{x}_0=\boldsymbol{DF}\boldsymbol{x}_0=(\boldsymbol{DF})^2\boldsymbol{x}_0$. 同理可得$\boldsymbol{x}_0=(\boldsymbol{DF})^N\boldsymbol{x}_0$. 因$(\boldsymbol{DF})^N=\boldsymbol{0}$，则必须$\boldsymbol{x}_0=\boldsymbol{0}$，与$\boldsymbol{I}-\boldsymbol{DF}$不是正则矩阵相矛盾. 因此，$\boldsymbol{I}-\boldsymbol{DF}$是正则矩阵，由定理 3 可以证明该系成立。（证明完毕）

下面，我们来考察由方框图得出的有向图，系 1 条件的具体意义. 该图是，将方框的输入输出换成节点，其连接关系用枝表示的图. 具体画法如下. 首先，对应于$S_i(i=1, 2, \cdots, P)$的输入画节点i，对应输出画节点$P+i$. 此外，若S_i具有直接通路，则画出从i* 到$P+i$的枝；若S_j的输出接到S_i的输入，则画出从$P+j$到i** 的枝. 象这样画出的图，表示方框之间及方框内部的静态连接关系. 容易看到，其相伴矩阵为

$$\begin{bmatrix} \boldsymbol{0} & \boldsymbol{F}^* \\ \boldsymbol{D}^* & \boldsymbol{0} \end{bmatrix}.$$

1) 注意，其逆未必成立.

* 原文误为$P+i$.——译者注

** 原文误为$P+i$.——译者注

如前面所述，因该相伴矩阵的 $2(N+1)$ 次幂可以表示成

$$\begin{bmatrix} \mathbf{0} & \boldsymbol{F}^* \\ \boldsymbol{D}^* & \mathbf{0} \end{bmatrix}^{2(N+1)} = \begin{bmatrix} \boldsymbol{F}^*(\boldsymbol{D}^*\boldsymbol{F}^*)^N\boldsymbol{D}^* & \mathbf{0} \\ \mathbf{0} & (\boldsymbol{D}^*\boldsymbol{F}^*)^{N+1} \end{bmatrix},$$

则系 1 的条件，即有使 $(\boldsymbol{D}^*\boldsymbol{F}^*)^N=\mathbf{0}$ 的正整数 N 存在意味着该图没有回路存在（参看相伴矩阵的性质(f)）．因此可以说，系 1 的条件表示方框图中没有静态回路．

由以上分析立即可见，满足系 1 条件具体的方框图，有下列几种情况：

(i) 当所有的方框都没有直接通路，即 $\boldsymbol{D}^*=\mathbf{0}$ 时；

(ii) 当方框之间不连接，即 $\boldsymbol{F}^*=\mathbf{0}$ 时；

(iii) 当整个系统为单向作用系统构造，即若适当改变方框编号，使得

$$\boldsymbol{F}^*=\begin{bmatrix} 0 & * & \cdots & * \\ 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & & * \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

时．式中 $*$ 表示是 1 或 0．

如上所述，能够以各方框状态变量的直积为状态变量，用标准形状态方程式和输出方程式完全表示整个系统的充分和必要条件是，$\det[\boldsymbol{I}-\boldsymbol{DF}]\neq 0$．下面我们再来讨论一下包括 $\boldsymbol{I}-\boldsymbol{DF}$ 不是正则在内的更一般的情况．在这种情况下，首先不知道表示整个系统的中间标准形式(9)是否有唯一解，而且既使有唯一解，由(9)式能得出那一种形式的状态方程式和输出方程式也是个问题．为了讨论这样一些问题，可以利用 A-I-2 中的定理 3．根据该定理，由中间标准形(9)式能够得到与其等价的（包含输入微分的）标准形状态方程式和输出方程式（因而(9)式具有唯一解）的充分和必要条件是，s 的有理多项式

$$D(s)\triangleq\det\begin{bmatrix} \boldsymbol{I} & -\boldsymbol{F} & \mathbf{0} & \mathbf{0} \\ -\boldsymbol{D} & \boldsymbol{I} & \mathbf{0} & -\boldsymbol{C} \\ \mathbf{0} & -\boldsymbol{J} & \boldsymbol{I} & \mathbf{0} \\ -\boldsymbol{B} & \mathbf{0} & \mathbf{0} & s\boldsymbol{I}-\boldsymbol{A} \end{bmatrix} \tag{15}$$

不恒等于零，即不是多项式环 $P(s,\ R)$ 的零元．利用 P-I-2(ix) 和 B-I-1 定理 1，可以计算出该行列式 $D(s)$

$$D(s)=D(s)\times 1$$

$$=\det\begin{bmatrix} \boldsymbol{I} & -\boldsymbol{F} & \mathbf{0} & \mathbf{0} \\ -\boldsymbol{D} & \boldsymbol{I} & \mathbf{0} & -\boldsymbol{C} \\ \mathbf{0} & -\boldsymbol{J} & \boldsymbol{I} & \mathbf{0} \\ -\boldsymbol{B} & \mathbf{0} & \mathbf{0} & s\boldsymbol{I}-\boldsymbol{A} \end{bmatrix}\det\begin{bmatrix} \boldsymbol{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \boldsymbol{C}(s\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}+\boldsymbol{D} & \boldsymbol{I} & \mathbf{0} & \mathbf{0} \\ \boldsymbol{J}(\boldsymbol{C}(s\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}+\boldsymbol{D}) & \boldsymbol{J} & \boldsymbol{I} & \mathbf{0} \\ (s\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B} & \mathbf{0} & \mathbf{0} & \boldsymbol{I} \end{bmatrix}$$

$$=\det\begin{bmatrix} \boldsymbol{I}-\boldsymbol{F}(\boldsymbol{C}(s\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}+\boldsymbol{D}) & -\boldsymbol{F} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I} & \mathbf{0} & -\boldsymbol{C} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & s\boldsymbol{I}-\boldsymbol{A} \end{bmatrix}$$

$$=\det[\boldsymbol{I}-\boldsymbol{F}(\boldsymbol{C}(s\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}+\boldsymbol{D})]\det[s\boldsymbol{I}-\boldsymbol{A}].$$

因 $\det[s\boldsymbol{I}-\boldsymbol{A}]$ 不是 $P(s,\ R)$ 的零元，则 $D(s)$ 不是 $P(s,\ R)$ 的零元和 $\det[\boldsymbol{I}-\boldsymbol{F}(\boldsymbol{C}(s\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}+\boldsymbol{D})]$ 不是有理体 $Q(s,\ R)$ 的零元等价．因此，得如下定理（参看 B-I-2 系 2）．

［定理 4］ 整个系统具有唯一解，且可用（包含输入微分的）标准形状态方程式及输出方程式完全表示的充分和必要条件是，

$$\det[\boldsymbol{I}-\boldsymbol{F}(\boldsymbol{C}(s\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}+\boldsymbol{D})] \tag{16}$$
$$(=\det[\boldsymbol{I}-\boldsymbol{C}(s\boldsymbol{I}-\boldsymbol{A})^{-1}(\boldsymbol{B}+\boldsymbol{D})\boldsymbol{F}])$$

不是有理函数体 $Q(s, R)$ 的零元．因 $\boldsymbol{A}$、$\boldsymbol{B}$、$\boldsymbol{C}$、$\boldsymbol{D}$ 均为分块对角形矩阵，则 $\boldsymbol{C}(s\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}+\boldsymbol{D}$ 也是分块对角形矩阵，即

$$\boldsymbol{C}(s\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}+\boldsymbol{D}=\begin{bmatrix} \boldsymbol{C}_1(s\boldsymbol{I}-\boldsymbol{A}_1)^{-1}\boldsymbol{B}_1+\boldsymbol{D}_1 & & \boldsymbol{0} \\ & \ddots & \\ \boldsymbol{0} & & \boldsymbol{C}_P(s\boldsymbol{I}-\boldsymbol{A}_P)^{-1}\boldsymbol{B}_P+\boldsymbol{D}_P \end{bmatrix}.$$

因该矩阵的 ii 子块是 S_i 的传递函数矩阵[31]，则定理 4 可以改述如下．

［系 2］ 整个系统具有唯一解，能用（包含输入微分的）标准形完全表示的充分和必要条件是，

$$\det[\boldsymbol{I}-\boldsymbol{F}\boldsymbol{H}(s)] \quad (=\det[\boldsymbol{I}-\boldsymbol{H}(s)\boldsymbol{F}]) \tag{17}$$

不是有理函数体 $Q(s, R)$ 的零元．式中 $\boldsymbol{H}(s)$ 是以各方框的传递函数矩阵 $\boldsymbol{H}_i(s)$ $(i=1, 2, \cdots, P)$ 为对角线子块的分块对角形矩阵

$$\boldsymbol{H}(s)=\begin{bmatrix} \boldsymbol{H}_1(s) & & \boldsymbol{0} \\ & \ddots & \\ \boldsymbol{0} & & \boldsymbol{H}_P(s) \end{bmatrix}. \tag{18}$$

（例 3） 现再来看一下例 2 中讨论过的系统（图 9）．因 s_1 的传递函数为 $(s+2)/(s+1)$，s_2 的传递函数为 $(s+4)/(s+3)$，则

$$\det[\boldsymbol{I}-\boldsymbol{F}\boldsymbol{H}(s)]=\det\begin{bmatrix} 1 & -\dfrac{s+4}{s+3} \\ -\dfrac{s+2}{s+1} & 1 \end{bmatrix}=-\frac{2s+5}{s^2+4s+3}.$$

因该有理函数不是 $Q(s, R)$ 的零元，根据系 2，可以保证图 9 所示系统能用标准形完全表示．因此，我们将按着 A-I-2 讲过的方法，由中间标准形推导出标准形．先将(12)，(13)，(14)式整理，写出(9)式的中间标准形

$$\left[\begin{array}{cc:cc:c:cc} 1 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ \hdashline -1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ \hdashline -1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 1 \end{array}\right]\begin{bmatrix} u_1 \\ u_2 \\ y_1 \\ y_2 \\ Z \\ \dot{x}_1 \\ \dot{x}_2 \end{bmatrix}$$

$$=\left[\begin{array}{cc} 0 & 0 \\ 0 & 0 \\ \hdashline 1 & 0 \\ 0 & 1 \\ \hdashline 0 & 0 \\ \hdashline -1 & 0 \\ 0 & -3 \end{array}\right]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}+\left[\begin{array}{c} 1 \\ 0 \\ \hdashline 0 \\ 0 \\ \hdashline 0 \\ \hdashline 0 \\ 0 \end{array}\right]V \tag{19}$$

对上式左边系数矩阵进行基本行变换，变成行标准形

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ y_1 \\ y_2 \\ Z \\ \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} -1 & 3 \\ 0 & 3 \\ 0 & 3 \\ -1 & 3 \\ -1 & 3 \\ -2 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \\ -1 \\ 0 \\ 1 \end{bmatrix} V \tag{20}$$

由该式可见，(19)式左边系数矩阵不是正则的．将式(20)的第7式展开，得

$$x_2 = -x_1 - V$$

将该关系代入其它6式，消去 x_2 和 $\dot{x}_2$，得

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ y_1 \\ y_2 \\ Z \\ \dot{x}_1 \end{bmatrix} = \begin{bmatrix} -4 \\ -3 \\ -3 \\ -4 \\ -4 \\ -5 \end{bmatrix} x_1 + \begin{bmatrix} -3 \\ -3 \\ -3 \\ -4 \\ -4 \\ -3 \end{bmatrix} V + \begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \\ -1 \\ -1 \end{bmatrix} \dot{V}. \tag{21}$$

显然，该式左边系数矩阵是正则的．因此，利用适当的基本行变换可以变成单位矩阵

$$\begin{bmatrix} u_1 \\ u_2 \\ y_1 \\ y_2 \\ Z \\ \dot{x}_1 \end{bmatrix} = \begin{bmatrix} -1.5 \\ -0.5 \\ -0.5 \\ -1.5 \\ -1.5 \\ -2.5 \end{bmatrix} x_1 + \begin{bmatrix} -1.5 \\ -1.5 \\ -1.5 \\ -2.5 \\ -2.5 \\ -1.5 \end{bmatrix} V + \begin{bmatrix} -0.5 \\ -0.5 \\ -0.5 \\ -0.5 \\ -0.5 \\ -0.5 \end{bmatrix} \dot{V} \tag{22}$$

取出该式下侧2式，便得到(包含输入微分的)标准形状态方程式和输出方程式

$$\dot{x}_1 = -2.5x_1 - 1.5V - 0.5\dot{V}$$
$$Z = -1.5x_1 - 2.5V - 0.5\dot{V}.$$

传递函数矩阵的推导方法

根据系2，整个系统可以用或许包含有输入微分的标准形状态方程式及输出方程式完全表示的充分和必要条件是 $\det[\boldsymbol{I}-\boldsymbol{F}\boldsymbol{H}(s)] \neq 0$．当该条件成立时，对于任意输入，整个系统有(由初始状态决定的)唯一的输出．因此，整个系统应该有传递函数矩阵存在，即下列结果成立．

［系3］ 当

$$\det[\boldsymbol{I}-\boldsymbol{F}\boldsymbol{H}(s)] \quad (=\det[\boldsymbol{I}-\boldsymbol{H}(s)\boldsymbol{F}]) \tag{17}$$

不是有理函数体 $Q(s, R)$ 的零元时，整个系统有传递函数矩阵存在．

那么，在这种情况下各方框的传递函数矩阵和整个系统的传递函数矩阵之间有什么关系呢？现整理一下各方框的传递函数矩阵，用(18)式 $\boldsymbol{H}(s)$ 的形式给出．设整个系统输入(的拉氏变换)为 $\boldsymbol{V}(s)$，输出为 $\boldsymbol{Z}(s)$，方框的输入为 $\boldsymbol{U}(s)$，输出为 $\boldsymbol{Y}(s)$，讨论其之间关

系．根据(5)′、(6)′、(8)′式，其拉氏变换形式为

$$\boldsymbol{U}(s)=\boldsymbol{F}\boldsymbol{Y}(s)+\boldsymbol{G}\boldsymbol{V}(s) \tag{23}$$

$$\boldsymbol{Z}(s)=\boldsymbol{J}\boldsymbol{Y}(s)+\boldsymbol{K}\boldsymbol{V}(s) \tag{24}$$

$$\boldsymbol{Y}(s)=\boldsymbol{H}(s)\boldsymbol{U}(s). \tag{25}$$

将(23)式代入(25)式，得

$$(\boldsymbol{I}-\boldsymbol{H}(s)\boldsymbol{F})\boldsymbol{Y}(s)=\boldsymbol{H}(s)\boldsymbol{G}\boldsymbol{V}(s)$$

因假定 $\det[\boldsymbol{I}-\boldsymbol{H}(s)\boldsymbol{F})\neq 0$，则可对 $\boldsymbol{Y}(s)$ 求解．由得到的 $\boldsymbol{Y}(s)$ 和(24)式可以求出整个系统的传递函数矩阵为

$$\boldsymbol{J}(\boldsymbol{I}-\boldsymbol{H}(s)\boldsymbol{F})^{-1}\boldsymbol{H}(s)\boldsymbol{G}+\boldsymbol{K}.$$

由下面的例子可见，系 3 是充分条件，然而其逆并不成立．这是由于传递函数矩阵仅表示输入输出之间的关系，它并不能表示系统内部不受输入影响部分及不受输出影响部分．

（例 4） 下面来讨论图 10 上方框图所表示的系统．整个系统的传递函数存在，是 1．此外，因

$$\boldsymbol{H}(s)=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix},\quad \boldsymbol{F}=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix},$$

则 $\det[\boldsymbol{I}-\boldsymbol{F}\boldsymbol{H}(s)]=0$．因此，系 3 的逆不成立．

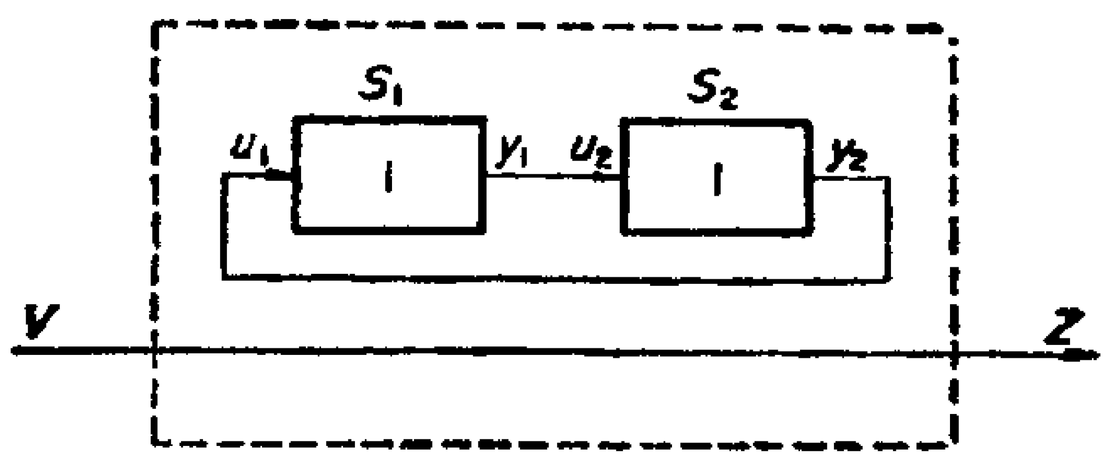


图 10 整个系统具有传递函数，而 $\det[\boldsymbol{I}-\boldsymbol{F}\boldsymbol{H}(s)]=0$ 的系统

B-I-4 克罗内克积，其它

本书开始讲过，矩阵可以看成是元素(实数、复数等等)的排列. 根据这个观点，若按一定规则改变某矩阵元素的排列，则可得出与原矩阵有关的新矩阵. 在 P-I-1 中，由 $\boldsymbol{A}=[a_{ij}]\in\tilde{R}^{m\times n}$ 作出 $\boldsymbol{A}^T=[a_{ji}]\in\tilde{R}^{n\times m}$ 就是其中最简单的一个例子.

本章将要叙述，元素按再稍为复杂一些的规律排列得到的矩阵的性质.

克罗内克(Kronecker)积

设 $\tilde{R}$ 为可换环，对于任意两个矩阵 $\boldsymbol{A}\in\tilde{R}^{m\times n}$ 和 $\boldsymbol{B}\in\tilde{R}^{p\times q}$，其克罗内克积或直积定义如下

$$\boldsymbol{A}\otimes\boldsymbol{B}\triangleq\begin{bmatrix} a_{11}\boldsymbol{B} & a_{12}\boldsymbol{B} & \cdots & a_{1n}\boldsymbol{B} \\ a_{21}\boldsymbol{B} & & & \vdots \\ \vdots & & & \vdots \\ a_{m1}\boldsymbol{B} & \cdots & \cdots & a_{mn}\boldsymbol{B} \end{bmatrix}\in\tilde{R}^{mp\times nq}, \tag{1}$$

即克罗内克积 $\boldsymbol{A}\otimes\boldsymbol{B}$ 是以 $a_{ij}\boldsymbol{B}$ 为 $i-j$ 子块的分块矩阵.

注意，克罗内克积与矩阵积(P-0-1)不同，它是对任意自然数组(m, n, p, q)定义的. 而且，显然 $\boldsymbol{I}_n\otimes\boldsymbol{I}_m=\boldsymbol{I}_{nm}$.

关于克罗内克积，下列基本公式成立

(a) 纯量倍

设 $\boldsymbol{A}$, $\boldsymbol{B}$ 为 $\tilde{R}$ 上的矩阵，$\alpha\in\tilde{R}$，则

$$\alpha(\boldsymbol{A}\otimes\boldsymbol{B})=(\alpha\boldsymbol{A})\otimes\boldsymbol{B}=\boldsymbol{A}\otimes(\alpha\boldsymbol{B}) \tag{2}$$

(b) 分配律

设 $\boldsymbol{A}$, $\boldsymbol{B}$ 是阶数相同的矩阵，则

$$(\boldsymbol{A}+\boldsymbol{B})\otimes\boldsymbol{C}=(\boldsymbol{A}\otimes\boldsymbol{C})+(\boldsymbol{B}\otimes\boldsymbol{C}) \tag{3}$$

$$\boldsymbol{C}\otimes(\boldsymbol{A}+\boldsymbol{B})=(\boldsymbol{C}\otimes\boldsymbol{A})+(\boldsymbol{C}\otimes\boldsymbol{B}) \tag{4}$$

(c) 结合律

$$\boldsymbol{A}\otimes(\boldsymbol{B}\otimes\boldsymbol{C})=(\boldsymbol{A}\otimes\boldsymbol{B})\otimes\boldsymbol{C} \tag{5}$$

(d) 转置

$$(\boldsymbol{A}\otimes\boldsymbol{B})^T=\boldsymbol{A}^T\otimes\boldsymbol{B}^T \tag{6}$$

(e) 混合积

设 $\boldsymbol{A}\in\tilde{R}^{k\times l}$, $\boldsymbol{B}\in\tilde{R}^{m\times n}$, $\boldsymbol{C}\in\tilde{R}^{p\times q}$, $\boldsymbol{D}\in\tilde{R}^{s\times t}$，则

$$\boldsymbol{A}\otimes\boldsymbol{C}=kp\times lq,\ \boldsymbol{B}\otimes\boldsymbol{D}=ms\times nt.$$

因此，若 $lq=ms$，则可定义$(\boldsymbol{A}\otimes\boldsymbol{C})(\boldsymbol{B}\otimes\boldsymbol{D})$. 而且，若

$$l=m,\ q=s,$$

则 $\boldsymbol{AB}$ 及 $\boldsymbol{CD}$ 这样的积可定义,

$$(\boldsymbol{A}\otimes\boldsymbol{C})(\boldsymbol{B}\otimes\boldsymbol{D})=\boldsymbol{AB}\otimes\boldsymbol{CD} \tag{7}$$

成立. 这样, AB、CD 可定义是 $(A\otimes C)(B\otimes D)$ 可定义的充分条件,但一般不是必要条件.

(f) 逆

对于 $A\in F^{n\times n}$、$B\in {}^{m\times m}$, 若有 A^{-1}、B^{-1} 存在, 则 $(A\otimes B)^{-1}$ 也存在, 且

$$(A\otimes B)^{-1}=A^{-1}\otimes B^{-1} \tag{8}$$

(g) 迹

对于 $A\in \widetilde{R}^{n\times n}$, $B\in \widetilde{R}^{m\times m}$, 有

$$\operatorname{tr}(A\otimes B)=\operatorname{tr}A\cdot \operatorname{tr}B. \tag{9}$$

在上列各式中, 因 (2) ~ (6) 式是显然的, 故证明省略, 这里只证明 (7) 和 (8) 式.

首先

$$(A\otimes C)(B\otimes D)=\{a_{ij}C\}\{b_{ln}D\}=\{\sum_k a_{ik}b_{kj}CD\}=\{(AB)_{ij}CD\}^{1)}=AB\otimes CD$$

利用 (7) 式, 由

$$(A\otimes B)(A^{-1}\otimes B^{-1})=(AA^{-1}\otimes BB^{-1})=(I_n\otimes I_m)=I_{nm}$$

得
$$(A\otimes B)^{-1}=A^{-1}\otimes B^{-1}.$$

上面证明是对于右逆矩阵, 显然对于左逆矩阵同样也可以证明. (8) 式表示, 两个矩阵的克罗内克积具有逆矩阵的充分条件是, 两个矩阵都存在着逆矩阵. 以后将会明白, 这同时也是必要条件.

[问题] 试证明 (9) 式.

扩大一下克罗内克积的概念, 我们定义克罗内克乘积

$$\underbrace{A\otimes A\otimes A\otimes\cdots\otimes A}_{k}=\dot{A}^k=n^k\times m^k \tag{10}$$

$$\dot{A}^0\triangleq 1\text{(纯数)} \tag{11}$$

$$\dot{A}^1\triangleq A \tag{12}$$

[例] 元素为 $+1$ 或 -1 的 $m\times m$ 矩阵 H, 当其满足

$$HH^T=mI_m$$

时, 称为 m 阶哈达马德 (Hadamard) 矩阵. 设 H_m 为 m 阶, H_n 为 n 阶哈达马德矩阵, 则 $H_m\otimes H_n$ 为 mn 阶哈达马德矩阵.

(证明)

$$\begin{aligned}(H_m\otimes H_n)(H_m\otimes H_n)^T&=(H_m\otimes H_n)(H_m^T\otimes H_n^T)=H_mH_m^T\otimes H_nH_n^T\\&=mI_m\otimes nI_n=mnI_{mn}.\end{aligned}$$

行展开及列展开

所谓矩阵 $A\in \widetilde{R}^{m\times n}$ 的行展开, 是将 A 的各行依次横排得到的 mn 维行向量, 列展开是将各列纵排得到的 mn 维列向量, 即

行展开

$$rsA\triangleq[a_{11}a_{12}\cdots a_{1n}a_{21}a_{22}\cdots a_{mn}] \tag{13}$$

列展开

$$csA\triangleq[a_{11}a_{21}\cdots a_{m1}a_{12}a_{22}\cdots a_{mn}]^T \tag{14}$$

---

1) $(AB)_{ij}$ 表示矩阵 AB 的第 i 行第 j 列上的元素.

(a) 转置

$$rsA^T=(csA)^T \tag{15}$$

$$csA^T=(csA)^T \tag{16}$$

(b) 积的展开

对于三个矩阵 $A\in\widetilde{R}^{m\times n}$, $B\in\widetilde{R}^{n\times p}$, $C\in\widetilde{R}^{p\times q}$, 有

$$rs(ABC)=rsB(A^T\otimes C) \tag{17}$$

$$cs(ABC)=(C^T\otimes A)csB \tag{18}$$

(证明) 现证明(17)式. 设 A 的第 i 行为 a_i, 则

$$rs(ABC)=[a_1BC,\ a_2BC,\ \cdots,\ a_mBC].$$

现仅取出 a_iBC 进行分析, 设 B 的第 j 行为 b_j, C 的第 k 列为 c_k, 则

$$a_iBC=a_i\begin{bmatrix} b_1c_1 & \cdots & b_1c_q \\ \vdots & & \vdots \\ b_nc_1 & \cdots & b_nc_q \end{bmatrix}=[b_1\ \cdots\ b_n]\begin{bmatrix} a_{i1}c_1 & \cdots & a_{i1}c_q \\ \vdots & & \vdots \\ a_{in}c_1 & \cdots & a_{in}c_q \end{bmatrix}$$

$$=rsB(a_i^T\otimes C)$$

$$\therefore\quad rs(ABC)=rsB[a_1^T\otimes C,\ a_2^T\otimes C,\ \cdots,\ a_m^T\otimes C]=rsB(A^T\otimes C).$$

(18)式的证明也是同样.

对于两个矩阵 $D\in\widetilde{R}^{s\times t}$, $E\in\widetilde{R}^{t\times u}$, 在(17)式中令 $m=n=s$, $p=t$, $q=u$, $A=I_s$, $B=D$, $C=E$, 则

$$rs(DE)=rsD(I_s\otimes E) \tag{19}$$

而且, 若在(17)式中令 $m=s$, $n=t$, $p=q=u$, $A=D$, $B=E$, $C=I_u$, 则

$$rs(DE)=rsE(D^T\otimes I_u) \tag{20}$$

归纳上列两式, 得

$$rs(DE)=rsD(I_s\otimes E)=rsE(D^T\otimes I_u) \tag{21}$$

[问题 1] 试证明两个对称埃尔米特矩阵的克罗内克积是对称埃尔米特矩阵.

[问题 2] 两个三角形矩阵的克罗内克积是三角形矩阵. 反之, 若 $A\otimes B$ 是不为 0 的三角形矩阵, 则 A 是三角形矩阵. 而且, 若 A 是正则矩阵, 则 B 也是三角形矩阵. 试证明之.

[问题 3] 若 A, B 均为对角矩阵, 则 $A\otimes B$ 也是对角矩阵. 反之, 若 $A\otimes B$ 是不为 0 的对角矩阵, 则 A, B 均为对角矩阵, 试证明之.

[问题 4] 试证明两个正交(酉)阵的克罗内克积是正交(酉)阵.

[问题 5] 对于 $A\in C^{n\times n}$, 一般将满足 $AA^*=A^*A$ 的矩阵称为正规阵 (normal matrix). 试证明两个正规阵的克罗内克积也是正规阵.

B-I-5 向量和矩阵的范数

我们知道，对于实数和复数，由于定义了它们的绝对值用来表示其大小，因而带来很多方便．例如，数列收敛的概念等也可以利用该绝对值定义．

因为收敛是分析中的基本概念，所以可以说实变函数及复变函数分析也是从引入绝对值的概念后开始的．

此外我们知道，解析几何及力学等当中讨论的普通意义的向量，也是根据上述绝对值的概念定义向量的大小．例如，平面向量 $\boldsymbol{x}$，当其在直角坐标中的分量为 x_1，x_2 时，可以用 $\sqrt{x_1^2+x_2^2}$ 给出其大小．

在 B-0-2 中，将向量的概念更一般化，讨论过向量空间的概念．对于平面（普通意义上的）向量，其大小可如上定义，而对于一般性向量，若引入在某种意义上表示其“大小”的纯量，对于讨论向量的收敛及分析也往往是很方便的．若将平面向量 $\boldsymbol{x}$ 的大小 $\sqrt{x_1^2+x_2^2}$ 写成 $\|\boldsymbol{x}\|$，显然该量具有和实数及复数的绝对值相似的性质，即具有下列性质：

（i） 若 $\boldsymbol{x}\neq\boldsymbol{\theta}$，则 $\|\boldsymbol{x}\|>0$，当 $\boldsymbol{x}=\boldsymbol{\theta}$ 时，则 $\|\boldsymbol{x}\|=0$；

（ii） $\|\alpha\boldsymbol{x}\|=|\alpha|\,\|\boldsymbol{x}\|$，$\alpha$ 为任意实数；

（iii） 对于两个向量 $\boldsymbol{x}_1$，$\boldsymbol{x}_2$，始终存在下列关系

$$\|\boldsymbol{x}_1+\boldsymbol{x}_2\|\leqslant\|\boldsymbol{x}_1\|+\|\boldsymbol{x}_2\|.\quad\text{（三角形不等式）}$$

对于表示向量大小的量，这些都是最基本的，对于一般性向量，再引入满足上述各性质的纯量（或者函数），用它来衡量向量的大小，并且称之为范数（norm）．

［定义 1］ 设体 F 仅为复数体 C 或实数体 R．对于体 F 上向量空间 $V(F)$ 中的向量 $\boldsymbol{x}$，$\boldsymbol{y}$，…，常把取实数值且具有下列性质的函数 $\|\cdot\|$ 称为向量的范数：

（i） 当 $\boldsymbol{x}\neq\boldsymbol{\theta}$ 时，$\|\boldsymbol{x}\|>0$，其中 $^\forall\boldsymbol{x}\in V$，$\|\boldsymbol{\theta}\|=0$；

（ii） $\|\alpha\boldsymbol{x}\|=|\alpha|\,\|\boldsymbol{x}\|$，其中 $^\forall\alpha\in F$，$^\forall\boldsymbol{x}\in V$；

（iii） $\|\boldsymbol{x}+\boldsymbol{y}\|\leqslant\|\boldsymbol{x}\|+\|\boldsymbol{y}\|$，其中 $^\forall\boldsymbol{x}$，$\boldsymbol{y}\in V$．

如下面几个例子所示，即使对于相同的 $V(F)$，也可以选取几个范数．

［例 1］ 在 R^n 或 C^n 上，令

$$\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix},$$

则可选取下列范数：

$$\|\boldsymbol{x}\|_1\triangleq|x_1|+|x_2|+\cdots+|x_n| \tag{1}$$

当 $p\geqslant1$ 固定为任意实数时

$$\|\boldsymbol{x}\|_p\triangleq[|x_1|^p+|x_2|^p+\cdots+|x_n|^p]^{\frac{1}{p}}_{\infty>p\geqslant1} \tag{2}$$

$$\|\boldsymbol{x}\|_\infty\triangleq\max_i|x_i|. \tag{3}$$

[例 2]　在有限区间$[t_0, t_1]$上定义的连续实函数的全体，构成 R 上的向量空间．由例 1 类推，对此可以考虑

$$\|f\|_1 \triangleq \int_{t_0}^{t_1} |f(t)|\,dt, \quad \|f\|_p \triangleq \left[\int_{t_0}^{t_1} |f(t)|^p\,dt\right]^{\frac{1}{p}}_{1 \leqslant p < \infty}$$

$$\|f\|_\infty = \max_{t \in [t_0, t_1]} |f(t)|$$

等等．

在 B-0-2 中讲过，体 F 上 $m \times n$ 矩阵的全体 $F^{m\times n}$ 形成向量空间．于是，当矩阵视为向量时，当然也可以定义范数．

[例 3]　设 $F=R$ 或 C，对于

$$\boldsymbol{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & & a_{mn} \end{bmatrix}, \ a_{ij} \in F$$

仿效[例 1]，当将 $\boldsymbol{A}$ 看作向量空间 $F^{m\times n}$ 中的向量时，可以考虑下列范数：

$$\|\boldsymbol{A}\|_{v_1} \triangleq \sum_{j=1}^{n}\sum_{i=1}^{m} |a_{ij}|$$

$$\|\boldsymbol{A}\|_{v_p} \triangleq \left[\sum_{j=1}^{n}\sum_{i=1}^{m} |a_{ij}|^p\right]^{\frac{1}{p}}, \ 1 \leqslant p < \infty$$

$$\|\boldsymbol{A}\|_{v_\infty} \triangleq \max_{i,j} |a_{ij}|.$$

(问题 1)

(ⅰ) 在 $C^n(R^n)$ 上，试证明下列关系

$$\|\boldsymbol{x}\|_1 \geqslant \|\boldsymbol{x}\|_2 \geqslant \|\boldsymbol{x}\|_\infty \geqslant \frac{1}{n}\|\boldsymbol{x}\|_1$$

(ⅱ) 试由此证明，$\|\boldsymbol{x}\|_1$，$\|\boldsymbol{x}\|_2$，$\|\boldsymbol{x}\|_\infty$ 互相等价[1]．亦即，若令其中任意两个范数为 $\|\cdot\|_a$，$\|\cdot\|_b$，则常有使

$$\alpha\|\boldsymbol{x}\|_a \leqslant \|\boldsymbol{x}\|_b \leqslant \beta\|\boldsymbol{x}\|_a, \ {}^\forall \boldsymbol{x} \in C^n$$

成立的常数 α，β 存在．

(ⅲ) 将以上结果推广到一般情况，试证明 $C^n(R^n)$ 上的任意范数都互相等价．

(问题 2)

在(2)式的 $\|\boldsymbol{x}\|_p$ 中，为何不令 $0<p<1$？对于 R^2 上的 2 维向量，试对 $p=1/2$，$p=1$，$p=2$，画出 $\|\boldsymbol{x}\|_p=1$ 的轨迹(即由范数 $\|\cdot\|_p$ 规定的"单位圆)，并比较之．

例 3 中矩阵的范数，当然具有[定义 1]中的性质(i)、(ii)、(iii)．但是，矩阵不仅具有向量的性质(即矩阵的和及纯量倍)，而且通常还要考虑矩阵的积．那么，对于矩阵的积上述范数具有什么性质呢？

设 $\boldsymbol{A}$，$\boldsymbol{B} \in F^{n\times n}(F=R$ 或 $C)$[2]

$$\begin{aligned}\|\boldsymbol{AB}\|_{v_1} &= \sum_{j=1}^{n}\sum_{i=1}^{n}\left|\sum_{k=1}^{n} a_{ik}b_{kj}\right| \leqslant \sum_{j=1}^{n}\sum_{i=1}^{n}\left\{\sum_{k=1}^{n}|a_{ik}|\sum_{k=1}^{n}|b_{kj}|\right\} \\ &= \left[\sum_{k=1}^{n}\sum_{i=1}^{n}|a_{ik}|\right]\left[\sum_{k=1}^{n}\sum_{j=1}^{n}|b_{kj}|\right] = \|\boldsymbol{A}\|_{v_1}\cdot\|\boldsymbol{B}\|_{v_1} \end{aligned} \tag{4}$$

---

[1] 当两个范数等价时，若对于其中一个范数向量收敛及有界，则对另外一个也必收敛及有界．

[2] 为了在 $\boldsymbol{A}$，$\boldsymbol{B} \in F^{m\times n}$ 中可以定义积 $\boldsymbol{AB}$，令 $m=n$．

同样

$$\|\boldsymbol{AB}\|_{v_2}\leqslant\|\boldsymbol{A}\|_{v_2}\cdot\|\boldsymbol{B}\|_{v_2}\tag{5}$$

成立.

但是,例如对于

$$\boldsymbol{A}=\begin{bmatrix}1&1\\0&1\end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix}1&0\\1&0\end{bmatrix}$$

显然$\|\boldsymbol{AB}\|_{v_\infty}\leqslant\|\boldsymbol{A}\|_{v_\infty}\|\boldsymbol{B}\|_{v_\infty}$并不成立[1]. 这往往不大方便,故在[定义 1]中再加上(4)、(5)式所示的性质后定义矩阵的范数.

[定义 2]

对于$\boldsymbol{A}\in F^{n\times n}$($F$ 是 R 或 C), 常把取实数值且具有下列性质的函数$\|\cdot\|$称为矩阵的范数.

(i) 当 $\boldsymbol{A}\neq\boldsymbol{0}_{n\times n}$ 时, $\|\boldsymbol{A}\|>0$, 其中 $^\forall\boldsymbol{A}\in F^{n\times n}$, $\|\boldsymbol{0}_{n\times n}\|=0$;

(ii) $\|\alpha\boldsymbol{A}\|=|\alpha|\,\|\boldsymbol{A}\|$, $^\forall\alpha\in F$, $^\forall\boldsymbol{A}\in F^{n\times n}$;

(iii) $\|\boldsymbol{A}+\boldsymbol{B}\|\leqslant\|\boldsymbol{A}\|+\|\boldsymbol{B}\|$, $^\forall\boldsymbol{A},\ \boldsymbol{B}\in F^{n\times n}$;

(iv) $\|\boldsymbol{AB}\|\leqslant\|\boldsymbol{A}\|\cdot\|\boldsymbol{B}\|$.

[例 4]

根据定义 2,

$$\|\boldsymbol{A}\|_{m_1}\triangleq\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|\tag{6}$$

$$\|\boldsymbol{A}\|_{m_2}\triangleq\left[\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2\right]^{1/2}\tag{7}$$

$$\|\boldsymbol{A}\|_{m_\infty}\triangleq n\max_{i,j}|a_{ij}|\tag{8}$$

均为矩阵 $\boldsymbol{A}\in F^{n\times n}$ 的范数.

(问题 3)

(i) 试证明$\|\boldsymbol{A}\|_{m_2}=[\mathrm{tr}\boldsymbol{A}^*\boldsymbol{A}]^{1/2}$.

(ii) 试证明$\|\boldsymbol{A}\|_{m_\infty}\triangleq n\max\limits_{i,j}|a_{ij}|$是矩阵的范数.

以上,考虑到矩阵的积, 在向量的性质中加上性质(iv)定义了矩阵的范数. 此外, 因为在矩阵的计算中, 矩阵往往和 F^n 中的向量混在一起, 故考虑矩阵和向量的积也是很重要的.

现在,我们对于 $\boldsymbol{x}\in F^n$, $\boldsymbol{A}\in F^{n\times n}$($F$ 是 R 或 C),考虑矩阵和向量的积 $\boldsymbol{Ax}$. 例如,对于[例 4]中的$\|\boldsymbol{A}\|_{m_1}$和推导(4)式的结果一样,可以证明

$$\|\boldsymbol{Ax}\|_1\leqslant\|\boldsymbol{A}\|_{m_1}\|\boldsymbol{x}\|_1.\tag{9}$$

一般,当

$$\|\boldsymbol{Ax}\|\leqslant\|\boldsymbol{A}\|\cdot\|\boldsymbol{x}\|,\ ^\forall\boldsymbol{x}\in F^n,\ ^\forall\boldsymbol{A}\in F^{n\times n}\tag{10}$$

成立时,称为矩阵范数$\|\boldsymbol{A}\|$和向量范数$\|\boldsymbol{x}\|$相容(compatible).

(问题 4)

试证明下列范数相容:

---

[1] 因此,$\|\boldsymbol{A}\|_{v_\infty}\triangleq\max\limits_{ij}|a_{ij}|$不能称为矩阵的范数.

(i) $\|\boldsymbol{A}\|_{m_2}$ 和 $\|\boldsymbol{x}\|_2$,

(ii) $\|\boldsymbol{A}\|_{m_\infty}$ 和 $\|\boldsymbol{x}\|_1$, $\|\boldsymbol{x}\|_2$.

一般,当向量 $\boldsymbol{x}\in F^n$ 的范数和矩阵 $\boldsymbol{A}\in F^{n\times n}$ 给定时,必然存在着和 $\|\boldsymbol{x}\|$ 相容的 $\boldsymbol{A}$ 的范数. 例如,可用下式确定这种矩阵的范数:

$$\|\boldsymbol{A}\|_i \triangleq \max_{x\neq 0}\frac{\|\boldsymbol{A}\boldsymbol{x}\|}{\|\boldsymbol{x}\|}\left(=\max_{\|x\|=1}\|\boldsymbol{A}\boldsymbol{x}\|\right) \tag{11}$$

因为这样求出来的矩阵范数 $\|\boldsymbol{A}\|$ 和向量的范数相容,故称为由向量范数导出(或诱导)的范数(induced norm). 实际上,当把矩阵看成线性作用子时,它是作用子范数. 关于这一点,在原讲座的第 II 部分将予以介绍.

[例 5]

$\|\boldsymbol{x}\|_1 \triangleq \sum_{i=1}^{n}|x_i|$ 的导出范数是

$$\|\boldsymbol{A}\|_{i_1} \triangleq \max_{j}\sum_{i=1}^{n}|a_{ij}| \tag{12}$$

$\|\boldsymbol{x}\|_2 \triangleq \left[\sum_{i=1}^{n}|x_i|^2\right]^{1/2}$ 的导出范数是

$$\|\boldsymbol{A}\|_{i_2} \triangleq \max_{i}\left[\lambda_i(\boldsymbol{A}^*\boldsymbol{A})\right]^{1/2} \tag{13}$$

式中 $\lambda_i(\boldsymbol{A}^*\boldsymbol{A})$ 是非负的埃尔米特阵 $\boldsymbol{A}^*\boldsymbol{A}$ 的第 i 个特征值.

$\|\boldsymbol{x}\|_\infty \triangleq \max_i|x_i|$ 的导出范数是

$$\|\boldsymbol{A}\|_{i_\infty} \triangleq \max_{i}\sum_{j=1}^{n}|a_{ij}| \tag{14}$$

显然,由向量范数导出的矩阵范数满足与原向量范数的相容性. 但是,关于定义 2 的性质(i)~(iv)又如何呢?

实际上可以证明,这些性质均成立,即

[定理 1]

设 $\boldsymbol{x}\in F^n$, $\boldsymbol{A}\in F^{n\times n}$($F$ 是 R 或 C), 由

$$\|\boldsymbol{A}\|_i \triangleq \max_{x=0}\frac{\|\boldsymbol{A}\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \tag{11}$$

定义的范数 $\|\boldsymbol{A}\|_i$, 具有定义 2 的性质(i)~(iv). 因此,也是在定义 2 意义上的矩阵范数. 而且,特别是对于 $\boldsymbol{A}=\boldsymbol{I}_n$, 始终是 $\|\boldsymbol{I}_n\|_i=1$.

定理 1 的证明及(12)、(13)、(14)式表示结果的正确性,例如在文献[24]中已详细讨论过. 由于不难,读者可试证之. 只是在证明(13)式时需要以后要讲的特征值的知识.

于是,对于矩阵范数,若使用由向量范数导出的矩阵范数,如上所述,由于它满足定义 2,而且和向量相容往往很方便,故以后主要使用这些矩阵范数. 而且,只写成 $\|\boldsymbol{A}\|$. 特别是,它表示由没有事先规定的向量范数导出的,我们将 $\|\boldsymbol{A}\|_{i_1}$ 等写成 $\|\boldsymbol{A}\|_1$.

(问题 5)

试证明 $\|\boldsymbol{A}\|\geqslant|a_{ij}|$, i, $j=1, 2, \cdots, n$, $^\forall\boldsymbol{A}\in\boldsymbol{C}^{n\times n}$.

(问题 6)

设 $\boldsymbol{A}\in C^{n\times n}$, $\det\boldsymbol{A}\neq 0$, 试证明

(i) $\|\boldsymbol{A}^{-1}\|\geqslant\|\boldsymbol{A}\|^{-1}$,

（ii）$\|A^{-1}\|^{-1}=\min\limits_{x\neq 0}\dfrac{\|Ax\|}{\|x\|}$.

（问题 7）

试证明，在 $F^{n\times n}$ 上，不仅由向量范数导出的，而且所有满足定义 2 的范数 $\|A\|$ 是互相等价的.（提示：所有满足定义 2 中条件(i)、(ii)、(iii)的都相互等价）

矩阵的测度[53]

和矩阵的范数一样，在系统理论中还有一个经常用到的量，称为矩阵的测度（measure）. 由于其定义与范数密切相关，所以在讨论微分方程式解的稳定性、大范围平衡点的存在性[53, 54]，网络数值解误差的评价[53, 55] 等时，利用测度比利用范数往往会得到更为明确的条件和评价. 下面先介绍测度的定义，之后再证明它的几个基本性质. 和讨论范数时一样，本节常假定体 F 是复数体 C 或实数体 R.

[定义 3]

矩阵 $A\in F^{n\times n}$ 的测度 $\mu(A)$ 是由下式确定的纯量（实数值）

$$\mu(A)\triangleq\lim_{h\to 0_+}\frac{\|I_n+hA\|-1}{h}\tag{15}$$

式中 $\|\cdot\|$ 是由向量范数导出的矩阵范数.

（注意）

（i）若注意到 $\|I_n\|=1$（参看定理 1），则意味着 (15) 式矩阵范数 $\|\cdot\|$ 在 I_n 点上沿 A 方向的方向微分.

（ii）由定义可见，即使对于同样的 $A\in F^{n\times n}$，当其范数 $\|A\|$ 不同（亦即，由于向量范数取得不同）时，$\mu(A)$ 值也会不同.

如下面例子所示，对于若干范数，利用定义式可以求出 $\mu(A)$ 的函数形式. 而且可以证明，一般对于任何向量范数往往都可以确定出 $\mu(A)$（参看附录）.

也就是说，和矩阵范数 $\|A\|$ 一样，当 A 给定，向量范数指定时，$\mu(A)$ 也常常是作为有限确定值确定的纯量.

[例 6]

对于向量范数 $\|x\|_1\triangleq\sum\limits_{i=1}^{n}|x_i|$,

$$\|A\|_1=\max_j\sum_{i=1}^{n}|a_{ij}|,$$

$$\mu_1(A)=\max_j\left\{\mathrm{Re}(a_{jj})+\sum_{\substack{i=1\\ i\neq j}}^{n}|a_{ij}|\right\}\tag{16}$$

对于向量范数 $\|x\|_2\triangleq\sqrt{\sum\limits_{i=1}^{n}|x_i|^2}$,

$$\|A\|_2=\max_i\sqrt{\lambda_i(A^*A)},$$

$$\mu_2(A)=\max_i\left\{\lambda_i\left(\frac{A^*+A}{2}\right)\right\}\tag{17}$$

对于向量范数 $\|x\|_\infty\triangleq\max\limits_i|x_i|$,

$$\|A\|_\infty=\max_i\sum_{j=1}^{n}|a_{ij}|,$$

$$\mu_\infty(\boldsymbol{A})=\max_i\left\{\mathrm{Re}(a_{ii})+\sum_{\substack{j=1\\j\neq i}}^{n}|a_{ij}|\right\} \tag{18}$$

这些结果可以推导如下.

首先

$$\begin{aligned}\|\boldsymbol{I}_n+h\boldsymbol{A}\|_1&=\max_j\sum_{i=1}^{n}|\delta_{ij}+ha_{ij}|\\&=\max_j\left\{|1+ha_{ij}|+\sum_{\substack{i=1\\i\neq j}}^{n}|ha_{ij}|\right\}\end{aligned}$$

(式中 $i=j$ 时, $\delta_{ij}=1$, $i\neq j$ 时, $\delta_{ij}=0$)

对于足够小的 $h>0$, 进而

$$=\max_j\left\{1+h\mathrm{Re}(a_{jj})+\circ(h)+\sum_{\substack{i=1\\i\neq j}}^{n}|ha_{ij}|\right\}$$

(式中 $\circ(h)$ 是当 $h\to0_+$ 时, $\circ(h)/h\to0$ 的项)

因此

$$\mu_1(\boldsymbol{A})\triangleq\lim_{h\to0_+}\{\|\boldsymbol{I}_n+h\boldsymbol{A}\|_1-1\}/h=\max_j\left\{\mathrm{Re}(a_{jj})+\sum_{\substack{i=1\\i\neq j}}^{n}|a_{ij}|\right\}.$$

对于 $\mu_\infty(\boldsymbol{A})$ 也是一样, 可以根据(15)式推导. 而且, 因为证明(14)式要用到后面将要讲到的矩阵特征值的性质, 所以这里只限于提出其结果.

下面, 我们来证明一下矩阵测度的基本性质.

表1 范数和测度的相似点

性质	范数 $\|\boldsymbol{A}\|$	测度 $\mu(\boldsymbol{A})$
齐次性	$\|\alpha\boldsymbol{A}\|=\|\alpha\|\,\|\boldsymbol{A}\|$, $\forall\alpha\in R$	$\mu(\alpha\boldsymbol{A})=\alpha\mu(\boldsymbol{A})$, $\alpha\geqslant0$
加法性	$\|\boldsymbol{A}+\boldsymbol{B}\|\leqslant\|\boldsymbol{A}\|+\|\boldsymbol{B}\|$	$\mu(\boldsymbol{A}+\boldsymbol{B})\leqslant\mu(\boldsymbol{A})+\mu(\boldsymbol{B})$
凸性	$\|\alpha\boldsymbol{A}+(1-\alpha)\boldsymbol{B}\|\leqslant\alpha\|\boldsymbol{A}\|+(1-\alpha)\|\boldsymbol{B}\|$, $0\leqslant\alpha\leqslant1$	$\mu(\alpha\boldsymbol{A}+(1-\alpha)\boldsymbol{B})\leqslant\alpha\mu(\boldsymbol{A})+(1-\alpha)\mu(\boldsymbol{B})$, $0\leqslant\alpha\leqslant1$
特征值的评价	$\mathrm{Re}[\lambda_i(\boldsymbol{A})]\leqslant\|\boldsymbol{A}\|$	$\mathrm{Re}[\lambda_i(\boldsymbol{A})]\leqslant\mu(\boldsymbol{A})$

$\mu(\boldsymbol{A})$的性质

(i) $\mu(\boldsymbol{I}_n)=1$, $\mu(-\boldsymbol{I}_n)=-1$ (19)

即必须注意, 它和范数不同, 由于 $\boldsymbol{A}$ 的形式不同, $\mu(\boldsymbol{A})$ 可能变成负的实数值.

(ii) $\mu(\boldsymbol{0}_{n\times n})=0$ (20)

但是 $\mu(\boldsymbol{A})=0$ 并不意味着 $\boldsymbol{A}=\boldsymbol{0}$.

(iii) $\mu(\alpha\boldsymbol{A})=\alpha\mu(\boldsymbol{A})$, $\forall\alpha\geqslant0$, $\forall\boldsymbol{A}\in F^{n\times n}$ (21)

$$\mu(\boldsymbol{A}+\alpha\boldsymbol{I}_n)=\mu(\boldsymbol{A})+\alpha,\ \forall\alpha\in R,\ \forall\boldsymbol{A}\in F^{n\times n}$$

第1个性质对应于范数 $\|\alpha\boldsymbol{A}\|=|\alpha|\,\|\boldsymbol{A}\|$, 不同的是有 $\alpha\geqslant0$ 的限制.

(iv) $-\|\boldsymbol{A}\|\leqslant-\mu(-\boldsymbol{A})\leqslant\mu(\boldsymbol{A})\leqslant\|\boldsymbol{A}\|$, $\forall\boldsymbol{A}\in F^{n\times n}$ (22)

这表示在范数和测度之间有 $|\mu(\boldsymbol{A})|\leqslant\|\boldsymbol{A}\|$ 的大小关系存在.

(v) $\mu(\boldsymbol{A}+\boldsymbol{B})\leqslant\mu(\boldsymbol{A})+\mu(\boldsymbol{B})$, $\forall\boldsymbol{A},\ \boldsymbol{B}\in F^{n\times n}$ (23)

该性质对应于范数$\|\boldsymbol{A}+\boldsymbol{B}\|\leqslant\|\boldsymbol{A}\|+\|\boldsymbol{B}\|$.

(vi) $-\mu(-\boldsymbol{A})\leqslant\mathrm{Re}\lambda_i(\boldsymbol{A})\leqslant\mu(\boldsymbol{A})$,

$$i=1,\ 2,\ \cdots,\ n,\ \forall\boldsymbol{A}\in F^{n\times n} \tag{24}$$

(vii) $\|\boldsymbol{Ax}\|\geqslant\max\{-\mu(-\boldsymbol{A}),\ -\mu(\boldsymbol{A})\}\cdot\|\boldsymbol{x}\|$,

$$\forall\boldsymbol{x}\in F^n,\ \forall\boldsymbol{A}\in F^{n\times n} \tag{25}$$

对应于范数的上限$\|\boldsymbol{Ax}\|\leqslant\|\boldsymbol{A}\|\cdot\|\boldsymbol{x}\|$,它给出下限.

(证明)

根据定义,(i)、(ii)、(iii)是显然的.

(iv) $\dfrac{\|\boldsymbol{I}_n-h\boldsymbol{A}\|-1}{h}+\dfrac{\|\boldsymbol{I}_n+h\boldsymbol{A}\|-1}{h}\geqslant\dfrac{\|2\boldsymbol{I}_n\|-2}{h}=0$

因此

$$-\|\boldsymbol{A}\|=\frac{-h\|\boldsymbol{A}\|+1-1}{h}\leqslant-\frac{\|\boldsymbol{I}_n-h\boldsymbol{A}\|-1}{h}\leqslant\frac{\|\boldsymbol{I}_n+h\boldsymbol{A}\|-1}{h}\leqslant\|\boldsymbol{A}\|$$

对于$h\to0_+$

$$-\|\boldsymbol{A}\|\leqslant-\mu(\boldsymbol{A})\leqslant\mu(\boldsymbol{A})\leqslant\|\boldsymbol{A}\|$$

(v) $\dfrac{\|\boldsymbol{I}_n+h(\boldsymbol{A}+\boldsymbol{B})\|-1}{h}=\dfrac{\|(\boldsymbol{I}_n+2h\boldsymbol{A})+(\boldsymbol{I}_n+2h\boldsymbol{B})\|-2}{2h}$

$$\leqslant\frac{\|\boldsymbol{I}_n+2h\boldsymbol{A}\|-1}{2h}+\frac{\|\boldsymbol{I}_n+2h\boldsymbol{B}\|-1}{2h}$$

因而

$$\mu(\boldsymbol{A}+\boldsymbol{B})\leqslant\mu(\boldsymbol{A})+\mu(\boldsymbol{B})$$

(vi) 虽然关于特征值以后才讲述,但是该性质的证明并不需要什么特征值的知识,故这里将予以证明. 设对应于$\boldsymbol{A}$的特征值λ_i的特征向量为$\boldsymbol{e}$(规格化成$\|\boldsymbol{e}\|=1$).

$$\frac{\|\boldsymbol{I}_n+h\boldsymbol{A}\|\|\boldsymbol{e}\|-1}{h}\geqslant\frac{\|\boldsymbol{e}+h\boldsymbol{A}\boldsymbol{e}\|-1}{h}=\frac{\|\boldsymbol{e}+h\lambda_i\boldsymbol{e}\|-1}{h}=\frac{|1+h\lambda_i|\,\|\boldsymbol{e}\|-1}{h}$$

$$=\frac{h\mathrm{Re}(\lambda_i)+\circ(h)}{h}$$

(设$h>0$为足够小)

由此得

$$\mu(\boldsymbol{A})\geqslant\mathrm{Re}(\lambda_i).$$

同样

$$-\mu(-\boldsymbol{A})\leqslant\mathrm{Re}(\lambda_i)$$

可以由

$$-\{\|\boldsymbol{I}_n-h\boldsymbol{A}\|-1\}/h\leqslant-\{|1-h\lambda_i|-1\}/h$$

加以证明.

(vii) $\|\boldsymbol{Ax}\|=\dfrac{\|(\boldsymbol{I}_n-h\boldsymbol{A})\boldsymbol{x}-\boldsymbol{x}\|}{h}\geqslant\dfrac{\|\boldsymbol{I}_n-h\boldsymbol{A}\|-1}{h}\|\boldsymbol{x}\|$

因此

$$\|\boldsymbol{Ax}\|\geqslant-\mu(-\boldsymbol{A})\|\boldsymbol{x}\|$$

而且,利用此结果得

$$\|\boldsymbol{Ax}\|=\|-\boldsymbol{Ax}\|\geqslant-\mu(\boldsymbol{A})\|\boldsymbol{x}\|.$$

于是可以看出,测度$\mu(\boldsymbol{A})$和范数$\|\boldsymbol{A}\|$一样,是取实数值的矩阵$\boldsymbol{A}$的函数,它和范数具有某些共同的性质. 当然,也存在一些重要的区别,如$\mu(\boldsymbol{A})$可能成为负值等等. 将其与$\|\boldsymbol{A}\|$相似的性质整理如下.

(问题 8) 试证明下列性质

(i) $\max\{\mu(\boldsymbol{A})-\mu(-\boldsymbol{B}),\ -\mu(\boldsymbol{A})+\mu(\boldsymbol{B})\}\leqslant\mu(\boldsymbol{A}+\boldsymbol{B})$

$$\forall\boldsymbol{A},\ \boldsymbol{B}\in F^{n\times n}$$

(ii) $|\mu(\boldsymbol{A})-\mu(\boldsymbol{B})|\leqslant\max\{|\mu(\boldsymbol{A}-\boldsymbol{B})|,\ |\mu(\boldsymbol{B}-\boldsymbol{A})|\}$

$$\forall\boldsymbol{A},\ \boldsymbol{B}\in F^{n\times n}$$

(问题 9) 试证明下列性质

(i) $\|\boldsymbol{A}^{-1}\|^{-1}\geqslant\max\{-\mu(-\boldsymbol{A}),\ -\mu(\boldsymbol{A})\}$,

$$\forall\boldsymbol{A}\in F^{n\times n},\ \det\boldsymbol{A}\neq 0$$

(ii) $\|(\boldsymbol{I}_n+h\boldsymbol{A})^{-1}\|^{-1}\geqslant 1-h\mu(-\boldsymbol{A})\geqslant 1-h\|\boldsymbol{A}\|$

$$\forall h\geqslant 0,\ \forall\boldsymbol{A}\in F^{n\times n}$$

(问题 10)

对于 $\boldsymbol{A}\in C^{n\times n}$, 当

$$\mathrm{Re}[a_{ii}]>\sum_{\substack{j=1\\j\neq i}}^{n}|a_{ij}|,\ i=1,\ 2,\ \cdots,\ n$$

成立时,称 $\boldsymbol{A}$ 是行和优势的(row-sum dominant).

同样,当

$$\mathrm{Re}[a_{jj}]>\sum_{\substack{i=1\\i\neq j}}^{n}|a_{ij}|,\ j=1,\ 2,\ \cdots,\ n$$

成立时,称 $\boldsymbol{A}$ 是列和优势的(column-sum dominant).

试证明, $\boldsymbol{A}$ 是行和优势(列和优势)的必要和充分和必要条件是

$$-\mu_{\infty}(-\boldsymbol{A})>(-\mu_1(-\boldsymbol{A})>0)$$

附录

在这里我们来证明,当 $\boldsymbol{A}\in C^{n\times n}$ 给定时, $\mu(\boldsymbol{A})$ 始终存在.

[定理][53~55]

对于由任意 $\boldsymbol{A}\in C^{n\times n}$ 和任意向量范数导出的矩阵范数, $\mu(\boldsymbol{A})$ 存在.

(证明)

对于 $0<\alpha<1$ 的实数 α,

$$\frac{\|\boldsymbol{I}_n+\alpha h\boldsymbol{A}\|-1}{\alpha h}=\frac{\|\alpha(\boldsymbol{I}_n+h\boldsymbol{A})+(1-\alpha)\boldsymbol{I}_n\|-1}{\alpha h}\leqslant\frac{\|\boldsymbol{I}_n+h\boldsymbol{A}\|-1}{h}$$

因此, 函数 $f(h)\triangleq\{\|\boldsymbol{I}_n+h\boldsymbol{A}\|-1\}/h$ 在 $h>0$ 处连续, 在 $h\to 0_+$ 时减小. 而且, $f(h)\geqslant-\|\boldsymbol{A}\|$(性质(iv))处有界. 因此, $\lim\limits_{h\to 0_+}f(h)\triangleq\mu(\boldsymbol{A})$.

B-I-6 向量和矩阵的收敛和极限

在 B-I-5 引入了向量和矩阵的范数这样的纯量，因为在某种意义上它相当于实数和复数的绝对值，所以也讲述了把向量和矩阵的分析结合起来的基本概念．下面，我们将利用范数来讨论向量和矩阵的收敛和极限．

在使用范数的关系上，本章也象 B-I-5 一样，所谈到的体 F，仅限于 $F=R$（实数体）或 $F=C$（复数体）．由范数的定义显而易见，这是因为对于纯量 $\alpha\in F$，其绝对值 $|\alpha|$ 是在普通意义上定义的．

［定义 1］

将体 F 上向量空间 $V(F)$ 的范数 $\|\cdot\|$ 固定．$V(F)$ 的向量序列 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_m, \cdots$ 用 $\{\boldsymbol{x}_m\}$ 表示．对于 $\{\boldsymbol{x}_m\}$，当存在某个向量 $\boldsymbol{x}\in V(F)$，使得 $\|\boldsymbol{x}_m-\boldsymbol{x}\|\to 0,\ m\to\infty$ 成立时，称为向量列序 $\{\boldsymbol{x}_m\}$ 收敛于 $\boldsymbol{x}$，$\boldsymbol{x}$ 称为该向量序列的极限．

（注意）

（i）当向量序列 $\{\boldsymbol{x}_m\}$ 收敛于 $\boldsymbol{x}$ 时，通常用 $\lim\limits_{m\to\infty}\boldsymbol{x}_m=\boldsymbol{x}$ 或 $\boldsymbol{x}_m\to\boldsymbol{x},\ m\to\infty$ 表示．

（ii）当给出向量序列 $\{\boldsymbol{x}_m\}$，判别其收敛性时，通常事先并不知道极限 $\boldsymbol{x}$（即使存在）．因此，不能直接利用定义 1 判断其收敛性．对于实数或复数数列的收敛性，利用所谓的柯西判别法是很方便的．此外，它也可以用于极限未知的情况，而且给出收敛的充分和必要条件．因此，在一般向量空间内，也可以考虑采用柯西判别法．一般，在范数 $\|\cdot\|$ 固定的向量空间 $V(F)$ 中[1)]，当 $n, m\to\infty$ 时，将 $\|\boldsymbol{x}_n-\boldsymbol{x}_m\|\to 0$ 的向量序列 $\{\boldsymbol{x}_m\}$ 称为柯西序列 (Cauchy Sequence)．那么是否象在 R 或 C 中一样，$\{\boldsymbol{x}_m\}$ 为柯西序列是 $\{\boldsymbol{x}_m\}$ 收敛的充分和必要条件呢？

实际上，它是必要条件（参看问题 1），但并不是充分条件[2)]．例如[28]，定义于 $t\in[0, 1]$ 上的实系数多项式的全体 $P[t, R]$，形成 R 上的向量空间．对于 $P[t, R]$ 上的向量序列和范数，现考虑

$$x_m(t)=\sum_{k=0}^{m}\frac{t^k}{k!},\quad \|x_m\|_\infty=\max_{t\in[0,1]}|x_m(t)|,$$

因 $\|x_n-x_m\|_\infty\leqslant\dfrac{1}{(m+1)!}+\cdots\dfrac{1}{n!}$，则 $\{x_m\}$ 是柯西序列．但是由于 $m\to\infty$ 时 $x_m(t)$ 趋近于 e^t，故 $P[t, R]$ 中没有极限，即该向量序列不收敛．

定义 1 给出一般向量空间中收敛的定义，若讨论范围仅限于应用中最重要的向量空间 F^n（即 R^n 或 C^n），情况又如何呢？

因此，现在我们来讨论任意向量序列 $\{\boldsymbol{x}_m\}$，$\boldsymbol{x}_m\in F^n$，设其具有极限 $\boldsymbol{x}\in F^n$．写成用分量表示的形式

1) 范数固定的向量空间 $V(F)$，称为（线性）赋范空间，表示成 $(V(F), \|\cdot\|)$ 等等．

2) 所有柯西序列都收敛于 $V(F)$ 中向量的空间 $(V(F), \|\cdot\|)$ 称为巴拿赫空间 (Banach Space)．

$$\boldsymbol{x}_m \triangleq \begin{bmatrix} x_1^{(m)} \\ x_2^{(m)} \\ \vdots \\ x_n^{(m)} \end{bmatrix} \qquad \boldsymbol{x} \triangleq \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{1}$$

暂且先假定使用下列范数

$$\|\boldsymbol{x}\|_\infty = \max_i |x_i| \tag{2}$$

由(1)、(2)两式立即可以证明

$$\|\boldsymbol{x}_m - \boldsymbol{x}\|_\infty = \max_i |x_i^{(m)} - x_i| \geqslant |x_i^{(m)} - x_i| \tag{3}$$

成立．显然，表示

$$\|\boldsymbol{x}_m - \boldsymbol{x}\|_\infty \to 0 \Rightarrow |x_i^{(m)} - x_i| \to 0,\ i=1,\ 2,\ \cdots,\ n \tag{4}$$

反之，若 $|x_i^{(m)} - x_i| \to 0$, $i=1,\ 2,\ \cdots,\ n$, 则 $\max\limits_i |x_i^{(m)} - x_i| = \|\boldsymbol{x}_m - \boldsymbol{x}\|_\infty \to 0$ 也成立．于是，得

$$\|\boldsymbol{x}_m - \boldsymbol{x}\|_\infty \to 0,\ m\to\infty \Leftrightarrow |x_i^{(m)} - x_i| \to 0,\ m\to\infty$$
$$(i=1,\ 2,\ \cdots,\ n) \tag{5}$$

但是，由于 F^n 上所有的范数都等价(B-I-5 中问题 1)，则对于任意范数，

$$\|\boldsymbol{x}_m - \boldsymbol{x}\|_\infty \to 0,\ m\to\infty \Leftrightarrow \|\boldsymbol{x}_m - \boldsymbol{x}\| \to 0,\ m\to\infty \tag{6}$$

成立．将(5)、(6)两式结合起来，最后得出下列结论："F^n 上向量序列 $\{\boldsymbol{x}_m\}$ 收敛于向量 $\boldsymbol{x}\in F^n$(对于任意范数)与 $\boldsymbol{x}_m$ 的各分量收敛于 $\boldsymbol{x}$ 的对应分量(在实数或复数数列的意义上)是一样的"．于是，对于 F^n 定义 1 可以叙述如下．

［定义 2］ 在向量空间 F^n 中，给定向量序列 $\{\boldsymbol{x}_m\}$，当 $\boldsymbol{x}_m$ 的各分量收敛时，称为 $\{\boldsymbol{x}_m\}$ 收敛．以各分量的极限为分量的向量 $\boldsymbol{x}\in F^n$ 称为 $\{\boldsymbol{x}_m\}$ 的极限．

(注意)

由于 F^n 中向量序列的收敛与各分量的收敛等价，所以可以使用 R(或 C)上的判别法判别其收敛性．例如，因为在 $R(C)$ 中形成柯西序列是收敛的充分和必要条件，所以在 F^n 中 $\{\boldsymbol{x}_m\}$ 为柯西序列也是收敛的充分和必要条件．

其次，我们来讨论矩阵序列的收敛性． 设给出 $F^{n\times n}$ 中的矩阵序列 $\boldsymbol{A}_1,\ \boldsymbol{A}_2,\ \cdots,\ \boldsymbol{A}_m,\ \cdots$，用 $\{\boldsymbol{A}_m\}$ 表示．因为矩阵由有限个元素组成，其收敛性可以和 F^n 中向量一样地考虑，所以我们也按照上述定义 1 或定义 2 规定矩阵的收敛性．

［定义 3］ 在 $F^{n\times n}$ 的矩阵序列 $\{\boldsymbol{A}_m\}$ 中，当各元素收敛时，称为 $\{\boldsymbol{A}_m\}$ 收敛． 以各元素的极限为元素的矩阵 $\boldsymbol{A}\in F^{n\times n}$ 称为 $\{\boldsymbol{A}_m\}$ 的极限．

它和 $\|\boldsymbol{A}_m - \boldsymbol{A}\| \to 0,\ m\to\infty$ 的条件等价． 这里对于矩阵范数，可以使用 B-I-5 中定义 2 给出的任意一个．

(注意)

$\{\boldsymbol{A}_m\}$ 收敛于 $\boldsymbol{A}$ 通常写成 $\lim\limits_{m\to\infty} \boldsymbol{A}_m = \boldsymbol{A}$ 或 $\boldsymbol{A}_m \to \boldsymbol{A},\ m\to\infty$.

(问题 1)

设向量空间 $V(F)$ 的向量 $\{\boldsymbol{x}_m\}$ 收敛于极限 $\boldsymbol{x}\in V(F)$，试证明这时 (i) $\{\boldsymbol{x}_m\}$ 是柯西序列；(ii) $\{\boldsymbol{x}_m\}$ 没有其它极限 $\boldsymbol{x}' \neq \boldsymbol{x}$.

(问题 2)

试证明，$F^{n\times n}$ 的矩阵序列 $\{\boldsymbol{A}_m\}$ 的各元素收敛于 $\boldsymbol{A}\in F^{n\times n}$ 的各元素与 $\|\boldsymbol{A}_m - \boldsymbol{A}\| \to 0$,

$m\to\infty$ 等价．并由此证明，$\{\boldsymbol{A}_m\}$ 收敛的充分和必要条件表示它是柯西序列．

（问题 3）

设 $\boldsymbol{A}\in F^{n\times n}$，试证明用 $\boldsymbol{A}$ 的幂作成的矩阵序列$\{\boldsymbol{A}^m\}$收敛于零阵 $\boldsymbol{0}_{n\times n}$ 的充分条件是，$\|\boldsymbol{A}\|<1$．（提示：$\|\boldsymbol{A}^m\|\leqslant\|\boldsymbol{A}\|^m$）

矩阵序列极限的性质

在 $F^{n\times n}$ 上给出矩阵序列 $\{\boldsymbol{A}_m\}$，$\{\boldsymbol{B}_m\}$，设其收敛于各自的极限 $\boldsymbol{A}\in F^{n\times n}$，$\boldsymbol{B}\in F^{n\times n}$，这时下列基本性质成立．

（i）和与差 $\{\boldsymbol{A}_m\pm\boldsymbol{B}_m\}$ 收敛

$$\lim_{m\to\infty}(\boldsymbol{A}_m+\boldsymbol{B}_m)=\boldsymbol{A}+\boldsymbol{B} \tag{7}$$

（ii）积 $\{\boldsymbol{A}_m\boldsymbol{B}_m\}$ 收敛

$$\lim_{m\to\infty}\boldsymbol{A}_m\boldsymbol{B}_m=\boldsymbol{A}\boldsymbol{B} \tag{8}$$

（iii）当逆矩阵 $\boldsymbol{A}_m^{-1}$，$\boldsymbol{A}^{-1}$ 均存在时，$\{\boldsymbol{A}_m^{-1}\}$ 收敛

$$\lim_{m\to\infty}\boldsymbol{A}_m^{-1}=\boldsymbol{A}^{-1} \tag{9}$$

（证明）因为 (i)、(ii) 很容易证明，故只证明 (iii)．$\boldsymbol{A}_m^{-1}=\operatorname{adj}\boldsymbol{A}_m/\det\boldsymbol{A}_m$．$\det\boldsymbol{A}_m$ 及 $\operatorname{adj}\boldsymbol{A}_m$ 的各元素都是 $\boldsymbol{A}_m$ 的元素的多项式．因为 $\lim\limits_{m\to\infty}$ 多项式($\boldsymbol{A}_m$ 的元素) = 多项式($\lim\limits_{m\to\infty}\boldsymbol{A}_m$ 的元素)，则 $\lim\limits_{m\to\infty}\operatorname{adj}\boldsymbol{A}_m=\operatorname{adj}\boldsymbol{A}$，$\lim\limits_{m\to\infty}\det\boldsymbol{A}_m=\det\boldsymbol{A}\neq 0$．因此，$\lim\limits_{m\to\infty}\boldsymbol{A}_m^{-1}=\operatorname{adj}\boldsymbol{A}/\det\boldsymbol{A}=\boldsymbol{A}^{-1}$．

（问题 4）

试证明当 $\lim\limits_{m\to\infty}\boldsymbol{A}_m=\boldsymbol{A}$，$\lim\limits_{m\to\infty}\boldsymbol{B}_m=\boldsymbol{B}$ 时，$\lim\limits_{m\to\infty}\boldsymbol{A}_m\otimes\boldsymbol{B}_m=\boldsymbol{A}\otimes\boldsymbol{B}$．

式中 $\otimes$ 表示克罗内克积(B-I-4)．

矩阵级数

下面我们来讨论矩阵级数．因为大家所熟悉的 e^A 等是作为无限级数的和定义的，所以在系统理论中有着重要应用．

[定义 4]

考察由 $F^{n\times n}$ 上矩阵序列 $\boldsymbol{A}_1$，$\boldsymbol{A}_2$，…，$\boldsymbol{A}_m$，… 组成的级数

$$\boldsymbol{A}_1+\boldsymbol{A}_2+\cdots+\boldsymbol{A}_m+\cdots \tag{10}$$

当用 $\boldsymbol{A}_m=[a_{ij}^{(m)}]$表示矩阵中的元素，由其元素作成的 R(或 C)的级数

$$a_{ij}^{(1)}+a_{ij}^{(2)}+\cdots+a_{ij}^{(m)}+\cdots \tag{11}$$

收敛时(对于所有的 i，$j=1, 2, \cdots, n$)，称为(10)式所示矩阵级数收敛，$\boldsymbol{A}=[a_{ij}]$ 称为级数的和．其中 a_{ij} 是(11)式的和．将矩阵用其各自的范数置换后得到的级数

$$\|\boldsymbol{A}_1\|+\|\boldsymbol{A}_2\|+\cdots+\|\boldsymbol{A}_m\|+\cdots \tag{12}$$

收敛时，称为(10)式所示矩阵级数绝对收敛．

（注意）

（i）矩阵级数 $\boldsymbol{A}_1+\boldsymbol{A}_2+\cdots+\boldsymbol{A}_m+\cdots$ 绝对收敛和元素的级数 $a_{ij}^{(1)}+a_{ij}^{(2)}+\cdots+a_{ij}^{(m)}$

$+\cdots$绝对收敛是等价的．因此，称为“绝对收敛”[1]．

（ii）利用矩阵序列的概念也可以表示矩阵级数的收敛性．亦即，若考察由部分和作成的矩阵序列 $\{\boldsymbol{S}_m\}$，$\boldsymbol{S}_m \triangleq \boldsymbol{A}_1+\boldsymbol{A}_2+\cdots+\boldsymbol{A}_m$．$\{\boldsymbol{S}_m\}$ 收敛于 $\boldsymbol{A}$ 与矩阵级数具有和 $\boldsymbol{A}$ 是等同的．

（例 1）

对于 $\boldsymbol{A}\in F^{n\times n}$，现来讨论下列级数

$$\boldsymbol{I}+\boldsymbol{A}+\frac{\boldsymbol{A}^2}{2!}+\cdots\frac{\boldsymbol{A}^k}{k!}+\cdots \tag{13}$$

（$\boldsymbol{I}$＝单位矩阵）

因

$$\|\boldsymbol{I}\|+\|\boldsymbol{A}\|+\cdots\frac{\|\boldsymbol{A}\|^k}{k!}+\cdots=e^{\|\boldsymbol{A}\|}-1+\|\boldsymbol{I}\|,$$

则该级数对于任意 $\boldsymbol{A}\in F^{n\times n}$ 绝对收敛．其和称为 $\boldsymbol{A}$ 的（矩阵）指数函数，用 $e^{\boldsymbol{A}}$ 表示．即

$$e^{\boldsymbol{A}}=\boldsymbol{I}+\boldsymbol{A}+\frac{\boldsymbol{A}^2}{2!}+\cdots\frac{\boldsymbol{A}^k}{k!}+\cdots \tag{14}$$

（例 2）

$$\cos\boldsymbol{A}=\boldsymbol{I}-\frac{\boldsymbol{A}^2}{2!}+\frac{\boldsymbol{A}^4}{4!}+\cdots(-1)^k\frac{\boldsymbol{A}^{2k}}{(2k)!}+\cdots \tag{15}$$

$$\sin\boldsymbol{A}=\boldsymbol{A}-\frac{\boldsymbol{A}^3}{3!}+\frac{\boldsymbol{A}^5}{5!}-\cdots(-1)^k\frac{\boldsymbol{A}^{2k+1}}{(2k+1)!}+\cdots \tag{16}$$

对于任意 $\boldsymbol{A}\in F^{n\times n}$ 都绝对收敛．

（例 3）

当 $\|\boldsymbol{A}\|<1$ 时，

$$\boldsymbol{I}+\boldsymbol{A}+\boldsymbol{A}^2+\cdots\boldsymbol{A}^k+\cdots \quad (\boldsymbol{A}\in F^{n\times n}) \tag{17}$$

绝对收敛．这时，和用 $\boldsymbol{S}$ 表示，我们来求一下 $\boldsymbol{S}$．设到 $(k+1)$ 项为止，$\boldsymbol{S}$ 的部分和为 $\boldsymbol{S}_k$，则恒等式

$$\boldsymbol{S}_k(\boldsymbol{I}-\boldsymbol{A})=\boldsymbol{I}-\boldsymbol{A}^{k+1} \tag{18}$$

成立．因 $\|\boldsymbol{A}\|<1$，则 $\det(\boldsymbol{I}-\boldsymbol{A})\neq 0$（B-I-5 中问题 7）．因此，由 (18) 式，得

$$\boldsymbol{S}_k=(\boldsymbol{I}-\boldsymbol{A})^{-1}-\boldsymbol{A}^{k+1}(\boldsymbol{I}-\boldsymbol{A})^{-1} \tag{19}$$

考虑到，对于 $k\to\infty$，$\boldsymbol{A}^{k+1}\to\boldsymbol{0}$．（参照问题 3）则

$$\lim_{k\to\infty}\boldsymbol{S}_k=\boldsymbol{S}=(\boldsymbol{I}-\boldsymbol{A})^{-1}. \tag{20}$$

即当 $\|\boldsymbol{A}\|<1$ 时，(17) 式无限矩阵级数与 $(\boldsymbol{I}-\boldsymbol{A})^{-1}$ 一致．

（注意）

例 3 证明了 $\|\boldsymbol{A}\|<1$ 是 $\boldsymbol{I}+\boldsymbol{A}+\cdots\boldsymbol{A}^k+\cdots$ 收敛的一个充分条件．我们知道，实际上

1）可证明如下．设 $\|\boldsymbol{A}_m\|_\infty\triangleq\max\limits_i\sum\limits_{j=1}^n|a_{ij}^{(m)}|$．显然，$\|\boldsymbol{A}_m\|_\infty\geqslant|a_{ij}^{(m)}|$，$i, j=1, 2, \cdots, n$．因此，若 $\sum\limits_{m=1}^\infty\|\boldsymbol{A}_m\|_\infty$ 收敛，则 $\sum\limits_{m=1}^\infty|a_{ij}^{(m)}|$ 也收敛．反之，当 $\sum\limits_{m=1}^\infty|a_{ij}^{(m)}|$ 收敛时，根据柯西判别法，对于任意的 $\varepsilon>0$，有满足

$$|a_{ij}^{(p+1)}|+|a_{ij}^{(p+2)}|+\cdots|a_{ij}^{(q)}|<\frac{\varepsilon}{n^2},\ \forall p,\ q\geqslant N(\varepsilon)\ (q>p)$$

的正整数 $N(\varepsilon)$ 存在．因此有 $\|\boldsymbol{A}_{p+1}\|_\infty+\|\boldsymbol{A}_{p+2}\|_\infty+\cdots\|\boldsymbol{A}_q\|_\infty<\varepsilon$．于是，再根据柯西判别法，$\sum\limits_{m=1}^\infty\|\boldsymbol{A}_\sigma\|_\infty$ 收敛．因为所有的矩阵范数是等价的，所以关于 $\|\cdot\|_\infty$ 得出的结论对于其它范数也正确．

$\lim\limits_{k\to\infty} \boldsymbol{A}^k=\boldsymbol{0}$(同样, 或者 $|\lambda_i(\boldsymbol{A})|<1$, $i=1, 2, \cdots, n$. λ_i 是 $\boldsymbol{A}$ 的特征值)是充分和必要条件[24]. 而且收敛时和与$(\boldsymbol{I}-\boldsymbol{A})^{-1}$一致.

此外例 3 也表明, $\boldsymbol{I}-\boldsymbol{A}$ 为正则矩阵的充分条件是 $\|\boldsymbol{A}\|<1$.

(问题 5)

试证明,在例 3 中只取部分和时的误差为[24]

$$\|(\boldsymbol{I}-\boldsymbol{A})^{-1}-\boldsymbol{S}_k\|\leqslant\|\boldsymbol{A}\|^{k+1}/(1-\|\boldsymbol{A}\|).$$

最后,关于两个绝对收敛矩阵级数,我们来证明下列应用起来很方便的结果.

[定理]

设 $F^{n\times n}$ 中的两个矩阵级数

$$S_1:\ \boldsymbol{A}_1+\boldsymbol{A}_2+\cdots\boldsymbol{A}_m+\cdots,\ S_2:\ \boldsymbol{B}_1+\boldsymbol{B}_2+\cdots\boldsymbol{B}_m+\cdots$$

都绝对收敛,其和分别为 $\boldsymbol{A}, \boldsymbol{B}$. 这时,将 S_1 和 S_2 按项相乘后作成的无限级数

$$\begin{aligned}S_3:\ &\boldsymbol{A}_1\boldsymbol{B}_1+(\boldsymbol{A}_1\boldsymbol{B}_2+\boldsymbol{A}_2\boldsymbol{B}_1)+(\boldsymbol{A}_1\boldsymbol{B}_3+\boldsymbol{A}_2\boldsymbol{B}_2+\boldsymbol{A}_3\boldsymbol{B}_1)+\cdots\\ &+(\boldsymbol{A}_1\boldsymbol{B}_m+\boldsymbol{A}_2\boldsymbol{B}_{m-1}+\cdots+\boldsymbol{A}_m\boldsymbol{B}_1)+\cdots\end{aligned}\tag{21}$$

绝对收敛,具有和 $\boldsymbol{AB}$.

(例 4)

对于

$$e^{\boldsymbol{A}}=\boldsymbol{I}+\boldsymbol{A}+\frac{\boldsymbol{A}^2}{2!}+\cdots,\ e^{\boldsymbol{B}}=\boldsymbol{I}+\boldsymbol{B}+\frac{\boldsymbol{B}^2}{2!}+\cdots,$$

应用上述定理,得

$$\begin{aligned}e^{\boldsymbol{A}}\cdot e^{\boldsymbol{B}}&=\left(\boldsymbol{I}+\boldsymbol{A}+\frac{\boldsymbol{A}^2}{2!}+\cdots\right)\left(\boldsymbol{I}+\boldsymbol{B}+\frac{\boldsymbol{B}^2}{2!}+\cdots\right).\\ &=\boldsymbol{I}+(\boldsymbol{A}+\boldsymbol{B})+\frac{1}{2!}(\boldsymbol{A}^2+\boldsymbol{AB}+\boldsymbol{BA}+\boldsymbol{B}^2)+\frac{1}{3!}(\boldsymbol{A}^3+3\boldsymbol{A}^2\boldsymbol{B}+3\boldsymbol{AB}^2+\boldsymbol{B}^3)+\cdots\end{aligned}\tag{22}$$

在这里,若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 可换,即 $\boldsymbol{AB}=\boldsymbol{BA}$, 则(22)式变成

$$e^{\boldsymbol{A}}\cdot e^{\boldsymbol{B}}=\boldsymbol{I}+(\boldsymbol{A}+\boldsymbol{B})+\frac{1}{2!}(\boldsymbol{A}+\boldsymbol{B})^2+\frac{1}{3!}(\boldsymbol{A}+\boldsymbol{B})^3+\cdots=e^{(\boldsymbol{A}+\boldsymbol{B})}.\tag{23}$$

(定理的证明)

根据假定,级数 $\sum\limits_{m=1}^{\infty}\|\boldsymbol{A}_m\|$, $\sum\limits_{m=1}^{\infty}\|\boldsymbol{B}_m\|$收敛. 对于绝对收敛的纯量级数, 按项相乘作成的级数也绝对收敛. 因此,级数

$$\|\boldsymbol{A}_1\|\cdot\|\boldsymbol{B}_1\|+(\|\boldsymbol{A}_1\|\cdot\|\boldsymbol{B}_2\|+\|\boldsymbol{A}_2\|\cdot\|\boldsymbol{B}_1\|)+\cdots\left(\sum_{i=1}^{m}\|\boldsymbol{A}_i\|\cdot\|\boldsymbol{B}_{m+1-i}\|\right)+\cdots\tag{24}$$

收敛. 将(24)和(21)的级数各项进行比较,显然 S_3 绝对收敛.

其次,设 S_1 和 S_2 到 m 项的部分和分别为 $\boldsymbol{S}_{1m}$、$\boldsymbol{S}_{2m}$, S_3 的部分和为 $\boldsymbol{S}_{3m}$. 并且,令级数 $\sum\limits_{m=1}^{\infty}\|\boldsymbol{A}_m\|$, $\sum\limits_{m=1}^{\infty}\|\boldsymbol{B}_m\|$的部分和分别为 $\bar{S}_{1m}$, $\bar{S}_{2m}$, 而级数(24)到m项的部分和为 $\bar{S}_{3m}$, 则

$$\begin{aligned}\boldsymbol{S}_{1m}\boldsymbol{S}_{2m}&=(\boldsymbol{A}_1\boldsymbol{B}_1+\boldsymbol{A}_1\boldsymbol{B}_2\cdots+\boldsymbol{A}_1\boldsymbol{B}_m)+\cdots(\boldsymbol{A}_m\boldsymbol{B}_1+\boldsymbol{A}_m\boldsymbol{B}_2+\cdots\boldsymbol{A}_m\boldsymbol{B}_m)\\ \boldsymbol{S}_{3m}&=(\boldsymbol{A}_1\boldsymbol{B}_1+\boldsymbol{A}_1\boldsymbol{B}_2+\cdots+\boldsymbol{A}_1\boldsymbol{B}_m)+\cdots+\boldsymbol{A}_m\boldsymbol{B}_1.\end{aligned}$$

因 $\bar{S}_{1m}\bar{S}_{2m}$, $\bar{S}_{3m}$ 分别是在上面两式中作置换 $\boldsymbol{A}_i\to\|\boldsymbol{A}_i\|$, $\boldsymbol{B}_i\to\|\boldsymbol{B}_i\|$ 后得到的,则

$$\|\boldsymbol{S}_{1m}\boldsymbol{S}_{2m}-\boldsymbol{S}_{3m}\|\leqslant\bar{S}_{1m}\bar{S}_{2m}-\bar{S}_{3m}.\tag{25}$$

在(25)式中, 令 $m\to\infty$, 则右边$\to 0$, 故得

$$\lim_{m\to\infty}\boldsymbol{S}_{1m}\boldsymbol{S}_{2m}=(\lim_{m\to\infty}\boldsymbol{S}_{1m})(\lim_{m\to\infty}\boldsymbol{S}_{2m})=\boldsymbol{AB}=\lim_{m\to\infty}\boldsymbol{S}_{3m}.$$

即证明 S_3 具有和 $\boldsymbol{AB}$.

(问题 6)

对于 $\boldsymbol{A}\in F^{n\times n}$, 设 $\sin\boldsymbol{A}$ 和 $\cos\boldsymbol{A}$ 由(16)、(15)式定义, 试证明下列结果.

(i) $\cos^2\boldsymbol{A}+\sin^2\boldsymbol{A}=\boldsymbol{I}$

(ii) 当 $\boldsymbol{AB}=\boldsymbol{BA}$ 时,

$$\sin(\boldsymbol{A}+\boldsymbol{B})=\sin\boldsymbol{A}\cos\boldsymbol{B}+\sin\boldsymbol{B}\cos\boldsymbol{A}$$

$$\cos(\boldsymbol{A}+\boldsymbol{B})=\cos\boldsymbol{A}\cos\boldsymbol{B}-\sin\boldsymbol{A}\sin\boldsymbol{B}.$$

$$e^{\boldsymbol{A}+j\boldsymbol{B}}=e^{\boldsymbol{A}}(\cos\boldsymbol{B}+j\sin\boldsymbol{B})$$

B-I-7 矩阵的微分

现在我们来讨论以实变数 x 的函数为元素的矩阵

$$\boldsymbol{A}(x)=\begin{bmatrix} a_{11}(x) & a_{12}(x) & \cdots & a_{1n}(x) \\ a_{21}(x) & a_{22}(x) & \cdots & a_{2n}(x) \\ \vdots & & & \vdots \\ a_{m1}(x) & a_{m2}(x) & \cdots & a_{mn}(x) \end{bmatrix} \tag{1}$$

式中设 a_{ij} 是 x 的实函数(或复变函数)，即对于任意 $x\in R$，设 $\boldsymbol{A}(x)\in R^{m\times n}$(或 $\boldsymbol{A}(x)\in C^{m\times n}$)．当 $a_{ij}(i=1,\ 2,\ \cdots,\ m,\ j=1,\ 2,\ \cdots,\ n)$ 可微分时，称为矩阵 $\boldsymbol{A}$ 可微．$\boldsymbol{A}$ 的各元素对 x 微分后，定义为 $\boldsymbol{A}$ 对 x 的微分[1)]，即

$$\frac{d\boldsymbol{A}(x)}{dx}(\triangleq \boldsymbol{A}'(x))\triangleq\left[\frac{da_{ij}}{dx}\right]. \tag{2}$$

(注意)

矩阵的微分也可以用和纯量函数情况下一样的形式定义．某个纯量函数(实函数或复变函数) $f(x)$ 的微分 $f'(x)$，基本上是一个满足下式的纯量：

$$\left|\frac{f(x+h)-f(x)}{h}-f'(x)\right|\to 0,\quad |h|\to 0 \tag{3}$$

将其用于矩阵的场合，用范数代替绝对值，则 $\boldsymbol{A}(x)$ 对 x 的微分可以定义成满足下式

$$\left\|\frac{\boldsymbol{A}(x+h)-\boldsymbol{A}(x)}{h}-\boldsymbol{A}'(x)\right\|\to 0,\quad |h|\to 0 \tag{4}$$

的矩阵 $\boldsymbol{A}'(x)\in F^{m\times n}$．因(4)式表示 $(\boldsymbol{A}(x+h)-\boldsymbol{A}(x))/h$ 的各元素(即 $(a_{ij}(x+h)-a_{ij}(x))/h$) 收敛于 $\boldsymbol{A}'(x)$ 的对应元素，则(4)式定义和(2)式定义是等价的(参看 B-1-6 中定义 3)．

显然，对于矩阵微分有下列基本公式．设 $\boldsymbol{A}$, $\boldsymbol{B}$ 均可微分，则

(i) 若 $\boldsymbol{A}$, $\boldsymbol{B}$ 是具有相同阶数的矩阵，则

$$\frac{d}{dx}(\boldsymbol{A}+\boldsymbol{B})=\frac{d\boldsymbol{A}}{dx}+\frac{d\boldsymbol{B}}{dx} \tag{5}$$

(ii) 在积 $\boldsymbol{AB}$ 可定义的情况下

$$\frac{d(\boldsymbol{AB})}{dx}=\left(\frac{d\boldsymbol{A}}{dx}\right)\boldsymbol{B}+\boldsymbol{A}\left(\frac{d\boldsymbol{B}}{dx}\right) \tag{6}$$

(iii) 若 a 是 x 的纯量函数，且对 x 可微分，则

$$\frac{d(a\boldsymbol{B})}{dx}=\left(\frac{da}{dx}\right)\boldsymbol{B}+a\left(\frac{d\boldsymbol{B}}{dx}\right) \tag{7}$$

(iv) 对于克罗内克积

1) 在这里，正确地讲，$\boldsymbol{A}'(x)$ 表示矩阵 $\boldsymbol{A}$ 的导数(derivative)．以下简称为微分．

$$\frac{d(\boldsymbol{A}\otimes\boldsymbol{B})}{dx}=\frac{d\boldsymbol{A}}{dx}\otimes\boldsymbol{B}+\boldsymbol{A}\otimes\frac{d\boldsymbol{B}}{dx} \tag{8}$$

(v) 若 A 为正则矩阵，则 $\boldsymbol{A}(x)^{-1}$ 亦可微分[9]

$$\frac{d(\boldsymbol{A}^{-1})}{dx}=-\boldsymbol{A}^{-1}\left(\frac{d\boldsymbol{A}}{dx}\right)\boldsymbol{A}^{-1}. \tag{9}$$

为引起注意，只证明(9)式.

$$\frac{d(\boldsymbol{A}^{-1}\boldsymbol{A})}{dx}=\boldsymbol{A}^{-1}\frac{d\boldsymbol{A}}{dx}+\left(\frac{d\boldsymbol{A}^{-1}}{dx}\right)\boldsymbol{A}=\frac{d\boldsymbol{I}}{dx}=\boldsymbol{0}$$

$$\therefore\quad \frac{d(\boldsymbol{A}^{-1})}{dx}=-\boldsymbol{A}^{-1}\left(\frac{d\boldsymbol{A}}{dx}\right)\boldsymbol{A}^{-1}.$$

由于(2)式定义的 $d\boldsymbol{A}/dx$ 仍然是一个矩阵，所以当 $a_{ij}(i=1,\ 2,\ \cdots,\ m;\ j=1,\ 2,\ \cdots,\ n)$ 两次可微分时，重复使用上述定义，无疑也可以定义二阶微分 $\frac{d^2\boldsymbol{A}}{dx^2}$. 而且，对于更高阶微分也是同样的.

其次，设某纯量 y 是向量 $\boldsymbol{x}$ 的各元素的函数，且对于各元素均可微分. 这时，y 对于 $\boldsymbol{x}$ 的微分变成叫做“梯度”(gradient)的向量，即

$$\frac{dy}{d\boldsymbol{x}}=\mathrm{grad}\,y. \tag{10}$$

grad y 是以 $\frac{\partial y}{\partial x_i}$ 为元素的向量. 它可以看作行向量，但有的文献上也把它看作列向量. 我们再看一下 y 对 $\boldsymbol{x}$ 的二阶微分 $d^2y/d\boldsymbol{x}^2$. 与其相应的是所谓的海赛矩阵 (Hessian Matrix)

$$\frac{d^2y}{d\boldsymbol{x}^2}=\left[\frac{\partial^2 y}{\partial x_i\partial x_j}\right]. \tag{11}$$

因为 $dy/d\boldsymbol{x}$ 是向量，所以向量对向量的微分一般可以看作是矩阵. 那么，$d^3y/d\boldsymbol{x}^3$，即矩阵对向量的微分又是什么呢？再进一步，矩阵对矩阵的微分又应该如何定义呢？

标量 y 对向量 $\boldsymbol{x}$ 的一阶微分是向量，即为元素的一维排列，二阶微分是矩阵，即为元素的二维排列. 由此可见，三阶微分可以定义成以 $\partial^3y/\partial\boldsymbol{x}_i\partial\boldsymbol{x}_j\partial\boldsymbol{x}_k$ 为$(i,\ j,\ k)$元素的三维立体排列. 但因立体排列不能表示在纸面上，所以必须按着适当的规则，变换成二维排列. 虽然变换规则很多，但是下面只介绍比较方便的一种，并说明其方便之处.

象前面一样，某矩阵 $\boldsymbol{A}=m\times n$ 对纯量 b_{kl} 的微分可以定义为

$$\frac{\partial\boldsymbol{A}}{\partial b_{kl}}\triangleq\left[\frac{\partial a_{ij}}{\partial b_{kl}}\right]=m\times n \tag{12}$$

现来考察以 $b_{kl}(k=1,\ 2,\ \cdots,\ p;\ l=1,\ 2,\ \cdots,\ q)$ 为$(k$、$l)$元素的矩阵 $\boldsymbol{B}=p\times q$. 当 $\boldsymbol{A}$ 对于 $\boldsymbol{B}$ 的各元素可微分时，称为 $\boldsymbol{A}$ 对 $\boldsymbol{B}$ 可微分. 将 $\boldsymbol{A}$ 对于 $\boldsymbol{B}$ 的微分定义成

$$\frac{d\boldsymbol{A}}{d\boldsymbol{B}}\triangleq\left[\frac{\partial\boldsymbol{A}}{\partial b_{kl}}\right]=\begin{bmatrix}\frac{\partial\boldsymbol{A}}{\partial b_{11}} & \frac{\partial\boldsymbol{A}}{\partial b_{12}} & \cdots & \frac{\partial\boldsymbol{A}}{\partial b_{1q}}\\ \vdots & & & \\ \frac{\partial\boldsymbol{A}}{\partial b_{p1}} & \cdots & \cdots & \frac{\partial\boldsymbol{A}}{\partial b_{pq}}\end{bmatrix}=mp\times nq \tag{13}$$

根据(13)式，定义下列运算子矩阵

$$\nabla_B=\left[\frac{\partial}{\partial b_{kl}}\right]=p\times q,$$

上述微分又可写成

$$\frac{d\boldsymbol{A}}{d\boldsymbol{B}}\triangleq\nabla_B\otimes\boldsymbol{A}.$$

因为 ∇_B 是以运算子为元素的矩阵,所以对于其中所包含的运算,克罗内克积的性质 (2)～(8) (B-I-4) 一般不成立. 因为 $d\boldsymbol{A}/d\boldsymbol{B}$ 是一个矩阵, 所以由此可以直接定义高阶微分. 即

$$\frac{d^2\boldsymbol{A}}{d\boldsymbol{B}^2}=\frac{d}{d\boldsymbol{B}}\left(\frac{d\boldsymbol{A}}{d\boldsymbol{B}}\right) \tag{14}$$

$$\frac{d^k\boldsymbol{A}}{d\boldsymbol{B}^k}=\frac{d}{d\boldsymbol{B}}\left(\frac{d^{k-1}\boldsymbol{A}}{d\boldsymbol{B}^{k-1}}\right) \tag{15}$$

等等.

由上面的定义, 前述问题一目了然. 即若 $\boldsymbol{y}\in R^m$, $\boldsymbol{x}\in R^n$(或 $\boldsymbol{y}\in C^m$, $\boldsymbol{x}\in C^n$) 皆为列向量, 则立即可以证明

$$\frac{d\boldsymbol{y}}{d\boldsymbol{x}^T}=\begin{bmatrix}\frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n}\\ \vdots & \vdots & & \vdots\\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_n}\end{bmatrix} \tag{16}$$

$$\frac{d\boldsymbol{y}^T}{d\boldsymbol{x}}=\left(\frac{d\boldsymbol{y}}{d\boldsymbol{x}^T}\right)^T \tag{17}$$

$$\frac{d\boldsymbol{y}}{d\boldsymbol{x}}=\operatorname{cs}\left[\frac{d\boldsymbol{y}}{d\boldsymbol{x}^T}\right] \tag{18}$$

$$\frac{d\boldsymbol{y}^T}{d\boldsymbol{x}^T}=\operatorname{rs}\left(\frac{d\boldsymbol{y}^T}{d\boldsymbol{x}}\right)=\operatorname{rs}\left[\left(\frac{d\boldsymbol{y}}{d\boldsymbol{x}^T}\right)^T\right]=\left[\operatorname{cs}\left(\frac{d\boldsymbol{y}}{d\boldsymbol{x}^T}\right)\right]^T \tag{19}$$

等等. 而且, 关于向量 $\boldsymbol{x}$, 有下列很有用的关系

$$\frac{d\boldsymbol{x}}{d\boldsymbol{x}^T}=\frac{d\boldsymbol{x}^T}{d\boldsymbol{x}}=\boldsymbol{I}_n$$

$$\frac{d\boldsymbol{x}}{d\boldsymbol{x}}=\operatorname{cs}\boldsymbol{I}_n$$

$$\frac{d\boldsymbol{x}^T}{d\boldsymbol{x}^T}=\operatorname{rs}\boldsymbol{I}_n. \tag{20}$$

关于矩阵 $\boldsymbol{A}$ 的 $\frac{d\boldsymbol{A}}{d\boldsymbol{A}}$, $\frac{d\boldsymbol{A}^T}{d\boldsymbol{A}}$, $\frac{d\boldsymbol{A}}{d\boldsymbol{A}^T}$ 等, 请参阅文献[56]. 此外, 根据定义立即可得出

$$\left(\frac{d\boldsymbol{A}}{d\boldsymbol{B}}\right)^T=\frac{d\boldsymbol{A}^T}{d\boldsymbol{B}^T}. \tag{21}$$

这些性质, 即使对于 $\partial\boldsymbol{B}/\partial\boldsymbol{A}$ 也可以同样证明.

矩阵积的微分

设某矩阵 $\boldsymbol{A}=n\times m$

$$\boldsymbol{A}=\boldsymbol{C}\boldsymbol{D}\quad \boldsymbol{C}=n\times r,\ \boldsymbol{D}=r\times m, \tag{22}$$

且 $\boldsymbol{C}$、$\boldsymbol{D}$ 均对 $\boldsymbol{B}=p\times q$ 可微分．下面来讨论 $d\boldsymbol{A}/d\boldsymbol{B}$ 将变成什么形式．首先，将 $\boldsymbol{A}$ 对纯量 b_{kl} 微分，显然

$$\frac{\partial \boldsymbol{A}}{\partial b_{kl}}=\frac{\partial}{\partial b_{kl}}(\boldsymbol{C}\boldsymbol{D})=\frac{\partial \boldsymbol{C}}{\partial b_{kl}}\boldsymbol{D}+\boldsymbol{C}\frac{\partial \boldsymbol{D}}{\partial b_{kl}}.$$

由此得

$$\frac{d\boldsymbol{A}}{d\boldsymbol{B}}=\left[\frac{\partial \boldsymbol{A}}{\partial b_{kl}}\right]=\begin{bmatrix}\dfrac{\partial \boldsymbol{C}}{\partial b_{11}}\boldsymbol{D} & \cdots & \dfrac{\partial \boldsymbol{C}}{\partial b_{1q}}\boldsymbol{D}\\ \vdots & & \vdots\\ \dfrac{\partial \boldsymbol{C}}{\partial b_{p1}}\boldsymbol{D} & \cdots & \dfrac{\partial \boldsymbol{C}}{\partial b_{pq}}\boldsymbol{D}\end{bmatrix}+\begin{bmatrix}\boldsymbol{C}\dfrac{\partial \boldsymbol{D}}{\partial b_{11}} & \cdots & \boldsymbol{C}\dfrac{\partial \boldsymbol{D}}{\partial b_{1q}}\\ \vdots & & \cdots\\ \boldsymbol{C}\dfrac{\partial \boldsymbol{D}}{\partial b_{p1}} & \cdots & \boldsymbol{C}\dfrac{\partial \boldsymbol{D}}{\partial b_{pq}}\end{bmatrix}$$

$$=\begin{bmatrix}\dfrac{\partial \boldsymbol{C}}{\partial b_{11}} & \cdots & \dfrac{\partial \boldsymbol{C}}{\partial b_{1q}}\\ \vdots & & \\ \dfrac{\partial \boldsymbol{C}}{\partial b_{p1}} & \cdots & \dfrac{\partial \boldsymbol{C}}{\partial b_{pq}}\end{bmatrix}\begin{bmatrix}\boldsymbol{D} & & \boldsymbol{0}\\ & \boldsymbol{D} & \\ \boldsymbol{0} & & \ddots\ \boldsymbol{D}\end{bmatrix}+\begin{bmatrix}\boldsymbol{C} & & \boldsymbol{0}\\ & \boldsymbol{C} & \\ \boldsymbol{0} & & \ddots\ \boldsymbol{C}\end{bmatrix}\begin{bmatrix}\dfrac{\partial \boldsymbol{D}}{\partial b_{11}} & \cdots & \dfrac{\partial \boldsymbol{D}}{\partial b_{1q}}\\ \vdots & & \vdots\\ \dfrac{\partial \boldsymbol{D}}{\partial b_{p1}} & \cdots & \dfrac{\partial \boldsymbol{D}}{\partial b_{pq}}\end{bmatrix}$$

$$=\frac{d\boldsymbol{C}}{d\boldsymbol{B}}(\boldsymbol{I}_q\otimes \boldsymbol{D})+(\boldsymbol{I}_p\otimes \boldsymbol{C})\frac{d\boldsymbol{D}}{d\boldsymbol{B}} \tag{23}$$

（问题）

当 $\boldsymbol{A}^{-1}$ 存在时，试证明 $\boldsymbol{A}^{-1}$ 对 $\boldsymbol{B}=p\times q$ 的微分

$$\frac{d\boldsymbol{A}^{-1}}{d\boldsymbol{B}}=-(\boldsymbol{I}_p\otimes \boldsymbol{A}^{-1})\frac{d\boldsymbol{A}}{d\boldsymbol{B}}(\boldsymbol{I}_q\otimes \boldsymbol{A}^{-1})$$

最后，我们再来看一下 $\boldsymbol{A}\otimes \boldsymbol{C}$ 对 $\boldsymbol{B}$ 的微分．和前面同样，因

$$\frac{\partial}{\partial b_{kl}}(\boldsymbol{A}\otimes \boldsymbol{C})=\frac{\partial \boldsymbol{A}}{\partial b_{kl}}\otimes \boldsymbol{C}+\boldsymbol{A}\otimes \frac{\partial \boldsymbol{C}}{\partial b_{kl}},$$

则

$$\frac{d}{d\boldsymbol{B}}(\boldsymbol{A}\otimes \boldsymbol{C})=\frac{d\boldsymbol{A}}{d\boldsymbol{B}}\otimes \boldsymbol{C}+\left[\boldsymbol{A}\otimes \frac{\partial \boldsymbol{C}}{\partial b_{kl}}\right] \tag{24}$$

必须注意，右边第二项不是 $\boldsymbol{A}\otimes \dfrac{d\boldsymbol{C}}{d\boldsymbol{B}}$.

B-I-8 矩阵的积分

我们来看以实变量 x 的函数为元素的矩阵 $\boldsymbol{A}(x)=[a_{ij}(x)]=m\times n$. 若 $a_{ij}(x)$ $(i=1, 2, \cdots, m;\ j=1, 2, \cdots, n)$对 x 可积,则以

$$\int_{x_1}^{x_2} a_{ij}(x)\, dx$$

为 $i-j$ 元素的矩阵,称为 $\boldsymbol{A}(x)$ 对 x 的积分,即

$$\int_{x_1}^{x_2} \boldsymbol{A}(x)\,dx \triangleq \left[\int_{x_1}^{x_2} a_{ij}(x)\,dx\right] \tag{1}$$

当然,不定积分也可以同样定义.

下面,对应于矩阵对矩阵的微分,我们来讨论对于矩阵的积分.作为它的预备知识,先来定义矩阵微分[1].

考虑矩阵 $\boldsymbol{A}=[a_{ij}]$的元素 a_{ij} 的微小变化 da_{ij},用 $d\boldsymbol{A}$ 表示以为 da_{ij} 为元素的矩阵,即

$$d\boldsymbol{A}\triangleq[da_{ij}]. \tag{2}$$

容易证明,和矩阵的微分 $d\boldsymbol{A}/dx$ 情况下相似的关系式

$$\begin{aligned}
&d(\boldsymbol{A}+\boldsymbol{B})=d\boldsymbol{A}+d\boldsymbol{B}\\
&d(\boldsymbol{A}\boldsymbol{B})=(d\boldsymbol{A})\boldsymbol{B}+\boldsymbol{A}(d\boldsymbol{B})\\
&d(a\boldsymbol{B})=(da)\boldsymbol{B}+a(d\boldsymbol{B})\\
&d(\boldsymbol{A}\otimes\boldsymbol{B})=(d\boldsymbol{A})\otimes\boldsymbol{B}+\boldsymbol{A}\otimes(d\boldsymbol{B})\\
&d(\boldsymbol{A}^{-1})=-\boldsymbol{A}^{-1}(d\boldsymbol{A})\boldsymbol{A}^{-1}
\end{aligned}$$

成立.

其次,我们来讨论 $\boldsymbol{A}=m\times n$ 的各元素是其它矩阵 $\boldsymbol{B}=p\times q$ 的元素 $b_{kl}(k=1, 2, \cdots, p,\ l=1, 2, \cdots, q)$的函数的情况.这时我们称 $\boldsymbol{A}$ 是 $\boldsymbol{B}$ 的函数,用 $\boldsymbol{A}(\boldsymbol{B})$ 表示.首先,因 $\boldsymbol{A}$ 的元素 a_{ij} 的微分

$$da_{ij}=\sum_{k=1}^{p}\sum_{l=1}^{q}\frac{\partial a_{ij}}{\partial b_{kl}}\,d\,b_{kl}=\frac{da_{ij}}{d(rs\,\boldsymbol{B})}\,d(rs\,\boldsymbol{B})^{T}=d(cs\,\boldsymbol{B})^{T}\frac{da_{ij}}{d(cs\,\boldsymbol{B})} \tag{3}$$

则

$$\begin{aligned}
d\boldsymbol{A}&=\begin{bmatrix}\dfrac{da_{11}}{d(rs\,\boldsymbol{B})}d(rs\,\boldsymbol{B})^{T} & \dfrac{da_{12}}{d(rs\,\boldsymbol{B})}d(rs\,\boldsymbol{B})^{T} & \cdots\\ \vdots & & \end{bmatrix}\\
&=\begin{bmatrix}\dfrac{da_{11}}{d(rs\,\boldsymbol{B})} & \dfrac{da_{12}}{d(rs\,\boldsymbol{B})} & \cdots\\ \vdots & & \end{bmatrix}\begin{bmatrix}d(rs\,\boldsymbol{B})^{T} & & \boldsymbol{0}\\ & \ddots & \\ \boldsymbol{0} & & d(\mathrm{rs}\,\boldsymbol{B})^{T}\end{bmatrix}\\
&=\frac{d\boldsymbol{A}}{d(rs\,\boldsymbol{B})}\{d(rs\,\boldsymbol{B})^{T}\otimes\boldsymbol{I}_n\}
\end{aligned} \tag{4}$$

1) 这里讲的微分是真正的微分(differential),必须和 B-I-7 中讲的微分(实际上指的导函数,即 derivative)区别开,因为一般不会混淆,故二者都称为微分.

同样

$$dA = \{d(\mathrm{cs}\,\boldsymbol{B})^T \otimes \boldsymbol{I}_n\} \frac{d\boldsymbol{A}}{d(\mathrm{cs}\,\boldsymbol{B})}. \tag{5}$$

于是，设某矩阵 $\boldsymbol{A}=m\times n$ 是矩阵 $\boldsymbol{B}=p\times q$ 的函数，$\boldsymbol{A}$ 对 $\boldsymbol{B}$ 的微分为 $d\boldsymbol{A}$. 这时，根据(4)或(5)式可以计算下列定积分

$$\int_{A_1}^{A_2} d\,\boldsymbol{A}(\boldsymbol{B})$$

即

$$\begin{aligned}\int_{\boldsymbol{A}_1}^{\boldsymbol{A}_2} d\boldsymbol{A}(\boldsymbol{B}) &= \int_{\boldsymbol{B}_1}^{\boldsymbol{B}_2} \frac{d\boldsymbol{A}}{d(\mathrm{rs}\,\boldsymbol{B})}\{d(\mathrm{rs}\,\boldsymbol{B})^T \otimes \boldsymbol{I}_n\} \\ &= \int_{\boldsymbol{B}_1}^{\boldsymbol{B}_2} \{d(\mathrm{cs}\,\boldsymbol{B})^T \otimes \boldsymbol{I}_n\} \frac{d\boldsymbol{A}}{d(\mathrm{cs}\,\boldsymbol{B})} \\ &= \left[\int_{\boldsymbol{B}_1}^{\boldsymbol{B}_2} \sum_{k=1}^{p}\sum_{l=1}^{q} \frac{\partial a_{ij}}{\partial b_{kl}}\, db_{kl}\right]\end{aligned} \tag{6}$$

该积分是线积分，其值与积分路线无关；而仅决定于 $\boldsymbol{B}_1$, $\boldsymbol{B}_2$[56].

A-I-5 由非线性方框组成的大系统

在 A-I-4 中，曾讨论过由几个用线性标准形方程式表示的子系统(方框)组成的动力学系统——大系统．研究了表示整个系统的方程式和表示各子系统的方程式之间的关系．讲过，即使整个系统由可以用标准形表示的子系统组成，但它也未必能用标准形表示等等．

这里我们将讨论子系统的状态方程式和输出方程式用非线性方程式 $\dot{\boldsymbol{x}}=\boldsymbol{f}(\boldsymbol{x},\ \boldsymbol{u})$，$\boldsymbol{y}=\boldsymbol{g}(\boldsymbol{x},\ \boldsymbol{u})$(以下为方便起见，称这两个方程式为正规形方程式)表示情况下的同样问题．首先是，在什么情况下整个系统能用和子系统相同形式的正规形方程式表示的问题．因为非线性系统的正规形包含线性标准形，后者是前者的特殊情况，所以不能说整个系统总可以用正规形表示．那么，在什么条件下才有可能呢，而且在这种条件下，与子系统的特性及结合方式的关系又如何？关于这些问题得出的结论，必须包括 A-I-4 中讲过的线性系统，因为后者是它的特殊情况．

其次，作为非线性系统的特殊问题，线性化中子系统与整个系统的关系是很重要的．如 A-I-1 中所述，以前在自动控制中可以作为线性系统处理的，多半是将非线性系统(在某个平衡状态附近)线性化后得到的系统．因此，在求整个系统的线性化表示式时就产生这样一个问题，是否能够认为，将子系统线性化得到的线性子系统组合起来构成的线性大系统就是上述的整个系统．

下面，将讨论这两个问题，并说明以前讲过的矩阵运算和微分如何应用．

系统的表示

和在 A-I-4 中一样，将所讨论的整个系统模型化，用方框图表示，假定该方框图由 P 个方框、若干相加点和引出点以及将它们连结起来的箭头所组成．设第 i 个方框 S_i 可以用下列正规形状态方程式和输出方程式

$$\dot{\boldsymbol{x}}_i=\boldsymbol{f}_i(\boldsymbol{x}_i,\ \boldsymbol{u}_i) \tag{1}$$

$$\boldsymbol{y}_i=\boldsymbol{g}(\boldsymbol{x}_i,\ \boldsymbol{u}_i) \tag{2}$$

表示．式中 $\boldsymbol{x}_i$ 是 n_i 维实列向量，表示状态；$\boldsymbol{u}_i$ 是 r_i 维实列向量，表示输入；$\boldsymbol{y}_i$ 是 m_i 维实列向量，表示输出．$\boldsymbol{f}_i$ 是由 $R^{n_i}\times R^{r_i}$ 到 R^{n_i} 的属于 C^1(连续可微函数类)的函数，$\boldsymbol{g}_i$ 是由 $R^{n_i}\times R^{r_i}$ 到 R^{m_i} 属于 C^1 的函数[1)]．

如在 A-I-4 中关于线性子系统讲过，放大器方框看成是状态向量维数为零($n_i=0$)的动力学子系统，不特别加以区别．为了使表示简单，下面将表示各方框的正规形(1)，(2)式归纳起来，写成

1) $\boldsymbol{f}_i$ 属于 C^1，也就是说，因为仅满足局部的利普希茨条件，所以该解不一定在起始时刻以后的全部时间内都存在．例如，$\dot{x}=x^2+v$ 对于 $\boldsymbol{v}=0$ 的解，根据给定的起始条件(即使是唯一的)，在有限时间上发散，即(1)式不一定满足状态方程式的条件(P. 3)．为了满足该条件，例如 $\boldsymbol{f}_i$ 必须满足大范围利普希茨条件，因为这里考察子系统的方程式的形式和整个系统方程式的形式之间的关系，故假定 $\boldsymbol{f}_i\in C^1$．

$$\dot{X}=f(X,\ U) \tag{3}$$

$$Y=g(X,\ U). \tag{4}$$

式中

$$X \triangleq \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix},\quad U \triangleq \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_p \end{bmatrix},\quad Y \triangleq \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix},$$

$$f(X,\ U) \triangleq \begin{bmatrix} f_1(x_1,\ u_1) \\ f_2(x_2,\ u_2) \\ \vdots \\ f_p(x_p,\ u_p) \end{bmatrix}$$

$$g(X,\ U) \triangleq \begin{bmatrix} g_1(x_1,\ u_1) \\ g_2(x_2,\ u_2) \\ \vdots \\ g_p(x_p,\ u_p) \end{bmatrix}.$$

各方框之间以及各方框与整个系统之间输入输出的连接关系，因与方框内部的特性完全无关，故可以利用 A-I-4 中讨论线性系统时定义的关联矩阵，用线性代数方程式

$$U=FY+GV \tag{5}$$

$$Z=JY+KV \tag{6}$$

表示．式中 V 和 Z 分别为 r 维和 m 维实列向量，表示整个系统的输入和输出．F 是表示各方框之间连接关系的 $\left(\sum_{i=1}^{P} r_i\right)\times\left(\sum_{i=1}^{P} m_i\right)$ 矩阵，G 是表示整个系统的输入和方框之间的连接关系的 $\left(\sum_{i=1}^{P} r_i\right)\times r$ 矩阵，J 是表示方框和整个系统输出之间连接关系的 $m\times\left(\sum_{i=1}^{P} m_i\right)$ 矩阵，K 是表示整个系统输入输出之间连接关系的 $m\times r$ 矩阵．

表示整个系统的正规形

整个系统可以用下列四个方程式

$$\dot{X}=f(X,\ U),\ f\in C' \tag{3}$$

$$Y=g(X,\ U),\ g\in C' \tag{4}$$

$$U=FY+GV \tag{5}$$

$$Z=JY+KV \tag{6}$$

表示．如 A-I-4 中所述，即使 f 和 g 是线性的，即(3)，(4)式是标准形 $\dot{X}=AX+BU$，$Y=CX+DU$，整个系统也未必具有唯一解，何况还不一定能用标准形 $\dot{X}=\tilde{A}X+\tilde{B}V$，$Z=\tilde{C}X+\tilde{D}V$ 表示．当然，这里所讨论的 f，g 是非线性函数的情况下，整个系统就更不一定具有唯一解．由(3)～(6)式消去 U 和 Y，不一定能得到下列正规形

$$\dot{X}=\tilde{f}(X,\ V),\ \tilde{f}\in C^1 \tag{7}$$

$$Z=\tilde{g}(X,\ V),\ \tilde{g}\in C^1. \tag{8}$$

那么，当那些条件成立时整个系统才具有唯一解，才能用正规形表示呢？

现在假定，由(3)—(6)式解出 U

$$\boldsymbol{U}=\boldsymbol{h}(\boldsymbol{X},\ \boldsymbol{V}),\ \boldsymbol{h}\in C^1. \tag{9}$$

则由(3)，(4)，(6)式得

$$\dot{\boldsymbol{X}}=\boldsymbol{f}(\boldsymbol{X},\ \boldsymbol{h}(\boldsymbol{X},\ \boldsymbol{V})) \tag{10}$$

$$\boldsymbol{Y}=\boldsymbol{g}(\boldsymbol{X},\ \boldsymbol{h}(\boldsymbol{X},\ \boldsymbol{V})) \tag{11}$$

$$\boldsymbol{Z}=\boldsymbol{J}\boldsymbol{g}(\boldsymbol{X},\ \boldsymbol{h}(\boldsymbol{X},\ \boldsymbol{V}))+\boldsymbol{K}\boldsymbol{V} \tag{12}$$

因 $\boldsymbol{f}$ 和 $\boldsymbol{h}$ 是连续可微的函数，则(10)式右边仍然连续可微，对于任意起始时刻 t_0 的起始条件 $\boldsymbol{X}=\boldsymbol{X}_0$ 和输入 $\boldsymbol{V}$，该式的解唯一存在．因此可以说，$\boldsymbol{U}$，$\boldsymbol{Y}$，$\boldsymbol{Z}$ 也由 $\boldsymbol{X}_0$ 和 $\boldsymbol{V}$ 唯一确定，整个系统具有唯一解，并且这时整个系统可以用正规形表示．

那么，我们来求 $\boldsymbol{U}$ 具有(9)式形式解的充分条件．将(4)式代入(5)式，消去 $\boldsymbol{Y}$，得

$$\boldsymbol{U}-\boldsymbol{F}\boldsymbol{g}(\boldsymbol{X},\ \boldsymbol{U})=\boldsymbol{G}\boldsymbol{V}. \tag{13}$$

利用大范围隐函数定理(Global Implicit Function Theorem)可以得到 $\boldsymbol{U}$ 具有(9)式形式解的充分条件．该定理如下所述．

［定理 1］[57]

设 $\hat{\boldsymbol{f}}$ 为从 $R^p\times R^q$ 到 R^p 的连续可微函数，即

$$\hat{\boldsymbol{f}}(\boldsymbol{x},\ \boldsymbol{u})=\boldsymbol{y}$$

$$\boldsymbol{x}\in R^p,\ \boldsymbol{u}\in R^q,\ \boldsymbol{y}\in R^p.$$

则解

$$\boldsymbol{x}=\hat{\boldsymbol{g}}(\boldsymbol{u},\ \boldsymbol{y}).$$

为由 $R^p\times R^q$ 到 R^p 连续可微的函数 $\hat{\boldsymbol{g}}$ 唯一存在的充分条件是

（ i ）对于所有的 $\boldsymbol{x}$ 和 $\boldsymbol{u}$，

$$\frac{\partial\hat{\boldsymbol{f}}}{\partial\boldsymbol{x}}\ \text{是正则的}$$

而且

（ii）对于所有的 $\boldsymbol{u}$，

$$\lim_{\|\boldsymbol{x}\|\to\infty}\|\hat{\boldsymbol{f}}(\boldsymbol{x},\ \boldsymbol{u})\|=\infty.$$

将该定理用于(13)式，可以得到 $\boldsymbol{U}$ 具有(9)式形式解的充分条件，亦即整个系统具有唯一解，可以用正规形表示的充分条件如下：

［定理 2］　若(i)对于所有的 $\boldsymbol{X}$ 和 $\boldsymbol{U}$，

$$\boldsymbol{I}-\boldsymbol{F}\frac{\partial\boldsymbol{g}}{\partial\boldsymbol{U}}\ \text{是正则的，}$$

及(ii)对于所有的 $\boldsymbol{X}$，

$$\lim_{\|\boldsymbol{U}\|\to\infty}\|\boldsymbol{U}-\boldsymbol{F}\boldsymbol{g}(\boldsymbol{X},\ \boldsymbol{U})\|=\infty,$$

则整个系统具有唯一解，可以用(7)，(8)式形式的正规形完全表示．

在各方框的输出方程式某些特殊形式下，该定理的条件(i)、(ii)非常简单．首先，当所有子系统输入输出之间没有直接通路时，

$$\boldsymbol{y}_i=\boldsymbol{g}_i(\boldsymbol{x}_i),\ \boldsymbol{g}_i\in C^1,\ i=1,\ 2,\ \cdots,\ P \tag{2$'$}$$

即可以表示成

$$\boldsymbol{Y}=\boldsymbol{g}(\boldsymbol{X}),\ \boldsymbol{g}\in C^1 \tag{4$'$}$$

因此

$$\boldsymbol{I}-\boldsymbol{F}\frac{\partial\boldsymbol{g}}{\partial\boldsymbol{U}}=\boldsymbol{I}$$

条件(i)总可以满足. 而且,因

$$\|U-Fg(X)\|\geqslant\|U\|-\|Fg(X)\|,$$

条件(ii)也始终能满足. 因此,得

[系 1] 若所有子系统输入输出之间没有直接通路,则整个系统具有唯一解,可以用(7),(8)式形式正规形完全表示.

现在,我们来讨论子系统输出方程式为下列形式

$$y_i=g_i(x_i)+D_iu_i,\ g_i\in C^1,\ i=1,\ 2,\ \cdots,\ P \tag{2)''$$

的情况. 将这些输出方程式归纳后,得

$$Y=g(X)+DU,\ g\in C^1 \tag{4)''$$

$$D\triangleq\begin{bmatrix} D_1 & 0 & \cdots & 0 \\ 0 & D_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & D_p \end{bmatrix}$$

于是,条件(i)、(ii)分别变成:

(i)' $I-FD$ 是正则的

(ii)' 对于所有的 X

$$\lim_{\|U\|\to\infty}\|(I-FD)U-Fg(X)\|=\infty.$$

但是,若 $I-FD$ 是正则的,因

$$\|(I-FD)U-Fg(X)\|\geqslant\|(I-FD)U\|-\|Fg(X)\|$$

$$\geqslant\frac{1}{\|(I-FD)^{-1}\|}\|U\|-\|Fg(X)\|$$

若条件(i)'满足,则条件(ii)总能满足(参看 B-I-5 中问题 6). 因此,得

[系 2] 当各子系统的输出方程式具有

$$y_i=g_i(x_i)+D_iu_i,\ g_i\in C^1,\ i=1,\ 2,\ \cdots,\ P$$

形式时,若

$$\det[I-FD]\neq 0,$$

$$D=\mathrm{diag}[D_1,\ D_2,\ \cdots,\ D_p],$$

则整个系统具有唯一解,可以用(7),(8)式形式正规形完全表示.

以上所得到的(13)式对 U 解成(9)式形式的充分条件,都是依据大范围隐函数定理.

对此,利用 A-I-4 中讲过的布尔矩阵进行分析,也可以得到系 1 扩大形式的另外的充分条件. 现将两个布尔矩阵 D^*, F^* 按下式定义:

$$D^*=\begin{bmatrix} d_{11}^* & 0 & \cdots\cdots & 0 \\ 0 & d_{22}^* & \cdots\cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots\cdots & d_{PP}^* \end{bmatrix}$$

$$d_{ii}^*\triangleq\begin{cases} 1: & \text{当 } S_i \text{ 中有直接通路,即输出方程式有 } y_i=g_i(x_i,\ u_i) \text{ 形式时} \\ 0: & \text{当 } S_i \text{ 中没有直接通路,即输出方程式可以写成 } y_i=g_i(x_i) \text{ 形式时} \end{cases}$$

$$i=1,\ 2,\ \cdots,\ P$$

$$\boldsymbol{F}^*=\begin{bmatrix} f_{11}^* & f_{12}^* & \cdots & f_{1P}^* \\ f_{21}^* & f_{22}^* & \cdots & f_{2P}^* \\ \vdots & \vdots & \ddots & \vdots \\ f_{P1}^* & f_{P2}^* & \cdots & f_{PP}^* \end{bmatrix}$$

$$f_{ij}^* \triangleq \begin{cases} 1; & \text{当 } S_j \text{ 的输出接到 } S_i \text{ 的输入, 即 } \boldsymbol{F}_{ij} \neq \boldsymbol{0} \text{ 时} \\ 0; & \text{当 } S_j \text{ 的输出不接到 } S_i \text{ 的输入, 即 } \boldsymbol{F}_{ij} = \boldsymbol{0} \text{ 时} \end{cases}$$

这时，若有满足 $(\boldsymbol{D}^*\boldsymbol{F}^*)^N=\boldsymbol{0}$ 的正整数 $N\leqslant P$ 存在，则所讨论的系统没有静态回路，和A-I-4 中线性情况一样．也就是说，若存在着这样的 N，则从任意一个方框的输入，沿着和连结各方框的箭头相反的方向和具有直接通路的方框，必然能够到达没有直连通路的方框或整个系统的输入，此外，没有其它通路．对此，可用数学式表示如下．由(13)式得

$$\boldsymbol{U}=\boldsymbol{Fg}(\boldsymbol{X},\ \boldsymbol{U})+\boldsymbol{GV}. \tag{13}'$$

将上列 $\boldsymbol{U}$ 的表示式反复代入上式右边，代入 N 次，得

$$\boldsymbol{U}=\overbrace{\boldsymbol{Fg}(\boldsymbol{X},\ \boldsymbol{Fg}(\boldsymbol{X},\ \boldsymbol{Fg}(\boldsymbol{X},\ \cdots\boldsymbol{Fg}(\boldsymbol{X},\ \boldsymbol{Fg}(\boldsymbol{X},\ \boldsymbol{U})}^{N+1}$$
$$+\overbrace{\boldsymbol{GV})+\boldsymbol{GV})\cdots+\boldsymbol{GV})+\boldsymbol{GV})+\boldsymbol{GV}}^{N+1}.$$

该式右边已经不是 $\boldsymbol{U}$ 的函数(与 $\boldsymbol{U}$ 无关)．而且，由该函数的形式可见，它是连续可微的函数，则(13)式可对 $\boldsymbol{U}$ 解成(9)式形式．故得下列定理．

[定理 3]

若有使

$$(\boldsymbol{D}^*\boldsymbol{F}^*)^N=\boldsymbol{0}$$

的正整数 $N\leqslant P$ 存在，则整个系统具有唯一解，可以用(7)，(8)式形式的正规形表示．

非线性大系统的线性化

前面已讲过，在控制理论中作为线性系统处理的，多半是将本来是非线性的系统在平衡状态附近线性化后得到的系统．现假定有一个非线性大系统，它由几个非线性方框构成．和以前一样，设各方框可以用正规形方程式

$$\dot{\boldsymbol{x}}_i=\boldsymbol{f}_i(\boldsymbol{x}_i,\ \boldsymbol{u}_i) \tag{1}$$
$$\boldsymbol{y}_i=\boldsymbol{g}_i(\boldsymbol{x}_i,\ \boldsymbol{u}_i) \tag{2}$$
$$i=1,\ 2,\ \cdots,\ P$$

表示．假定在该非线性系统上加上某个定值输入 $\boldsymbol{V}_0$ 时，各方框在固定状态 $\boldsymbol{X}=\boldsymbol{X}_0$ 处处于平衡状态，即假定有满足(3)～(6)式的常数向量解 $\boldsymbol{X}=\boldsymbol{X}_0$, $\boldsymbol{V}=\boldsymbol{V}_0$, $\boldsymbol{Z}=\boldsymbol{Z}_0$, $\boldsymbol{U}=\boldsymbol{U}_0$, $\boldsymbol{Y}=\boldsymbol{Y}_0$ 存在，使得下列各式

$$\boldsymbol{U}_0=\boldsymbol{FY}_0+\boldsymbol{GV}_0 \tag{14}$$
$$\boldsymbol{Z}_0=\boldsymbol{JY}_0+\boldsymbol{KV}_0 \tag{15}$$
$$\boldsymbol{0}=\boldsymbol{f}(\boldsymbol{X}_0,\ \boldsymbol{U}_0) \tag{16}$$
$$\boldsymbol{Y}_0=\boldsymbol{g}(\boldsymbol{X}_0,\ \boldsymbol{U}_0) \tag{17}$$

成立．

于是，若将各方框的方程式(1)，(2)在解 $\boldsymbol{X}_0$, $\boldsymbol{U}_0$, $\boldsymbol{Y}_0$ 附近线性化，便可得到 P 个线性标准形方程式．在原来的非线性方框图的每个方框中，将其正规形方程式(1)，(2)换成

刚才得到的标准形方程式，这样得到的线性大系统到底表示什么？直观地推测，它当然是原非线性大系统线性化后的系统．但是是否能够说，各子系统线性化后连接成的系统就是整个系统线性化后得到的系统？因此，关于这个问题下面再稍微讨论一下．

考虑下列问题：

$Q1$ "将各方框的特性在 $\boldsymbol{X}_0$, $\boldsymbol{U}_0$, $\boldsymbol{Y}_0$ 附近线性化，用所得到的标准形替换后的线性大系统表示什么，即它具有哪些与原非线性大系统有关的信息．这里假定，并不知道原非线性大系统是否可以用(7)，(8)式正规形方程式表示等等．"

以前，在线性控制理论中，当各子系统是以本来是非线性的系统为对象时，对于整个系统来讲，它是否具有正规形方程式等等，并不认为是个问题，一开始就是以各子系统线性化得到的系统组合成的线性系统为对象．对于小信号，它和原非线性系统的特性很接近，亦即，基于上述一些直观的假定，我们把对 $Q1$ 的讨论作为它的理论依据．

首先，将表示各子系统的正规形在

$$\begin{bmatrix} \boldsymbol{x}_{10} \\ \boldsymbol{x}_{20} \\ \vdots \\ \boldsymbol{x}_{P0} \end{bmatrix} \triangleq \boldsymbol{X}_0, \qquad \begin{bmatrix} \boldsymbol{u}_{10} \\ \boldsymbol{u}_{20} \\ \vdots \\ \boldsymbol{u}_{P0} \end{bmatrix} \triangleq \boldsymbol{U}_0, \qquad \begin{bmatrix} \boldsymbol{y}_{10} \\ \boldsymbol{y}_{20} \\ \vdots \\ \boldsymbol{y}_{P0} \end{bmatrix} \triangleq \boldsymbol{Y}_0$$

的附近线性化

$$\dot{\boldsymbol{x}}_{id} = \left[\frac{\partial \boldsymbol{f}_i}{\partial \boldsymbol{x}_i}\right]_{\substack{x_i = x_{i0} \\ u_i = u_{i0}}} \boldsymbol{x}_{id} + \left[\frac{\partial \boldsymbol{f}_i}{\partial \boldsymbol{u}_i}\right]_{\substack{x_i = x_{i0} \\ u_i = u_{i0}}} \boldsymbol{u}_{id} \tag{18}$$

$$\boldsymbol{y}_{id} = \left[\frac{\partial \boldsymbol{g}_i}{\partial \boldsymbol{x}_i}\right]_{\substack{x_i = x_{i0} \\ u_i = u_{i0}}} \boldsymbol{x}_{id} + \left[\frac{\partial \boldsymbol{g}_i}{\partial \boldsymbol{u}_i}\right]_{\substack{x_i = x_{i0} \\ u_i = u_{i0}}} \boldsymbol{u}_{id} \tag{19}$$

式中 $\boldsymbol{x}_{id}$, $\boldsymbol{u}_{id}$, $\boldsymbol{y}_{id}$ 分别表示 $\boldsymbol{x}_i$, $\boldsymbol{u}_i$, $\boldsymbol{y}_i$ 在 $\boldsymbol{x}_{i0}$, $\boldsymbol{u}_{i0}$, $\boldsymbol{y}_{i0}$ 附近的变化量．将原来的非线性子系统换成用这些标准形表示的线性化子系统，我们来讨论一下由此得到的线性大系统．和 A-I-4 中一样，将表示各子系统的标准形归纳起来，可以写成

$$\dot{\boldsymbol{X}}_d = \left[\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{X}}\right]_{\substack{X = X_0 \\ U = U_0}} \boldsymbol{X}_d + \left[\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{U}}\right]_{\substack{X = X_0 \\ U = U_0}} \boldsymbol{U}_d \tag{20}$$

$$\boldsymbol{Y}_d = \left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{X}}\right]_{\substack{X = X_0 \\ U = U_0}} \boldsymbol{X}_d + \left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X = X_0 \\ U = U_0}} \boldsymbol{U}_d \tag{21}$$

$$\boldsymbol{X}_d \triangleq \begin{bmatrix} \boldsymbol{x}_{1d} \\ \boldsymbol{x}_{2d} \\ \vdots \\ \boldsymbol{x}_{Pd} \end{bmatrix}, \qquad \boldsymbol{U}_d \triangleq \begin{bmatrix} \boldsymbol{u}_{1d} \\ \boldsymbol{u}_{2d} \\ \vdots \\ \boldsymbol{u}_{Pd} \end{bmatrix}, \qquad \boldsymbol{Y}_d \triangleq \begin{bmatrix} \boldsymbol{y}_{1d} \\ \boldsymbol{y}_{2d} \\ \vdots \\ \boldsymbol{y}_{Pd} \end{bmatrix}$$

其原因是，由 $\boldsymbol{f}$, $\boldsymbol{g}$ 的定义，

$$\left[\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{X}}\right]_{\substack{X = X_0 \\ U = U_0}} = \begin{bmatrix} \left[\frac{\partial \boldsymbol{f}_1}{\partial \boldsymbol{x}_1}\right]_{\substack{x_1 = x_{10} \\ u_1 = u_{10}}} & & & \boldsymbol{0} \\ & \left[\frac{\partial \boldsymbol{f}_2}{\partial \boldsymbol{x}_2}\right]_{\substack{x_2 = x_{20} \\ u_2 = u_{20}}} & & \\ & & \ddots & \\ \boldsymbol{0} & & & \left[\frac{\partial \boldsymbol{f}_P}{\partial \boldsymbol{x}_P}\right]_{\substack{x_P = x_{P0} \\ u_P = u_{P0}}} \end{bmatrix} \tag{22}$$

$$\left[\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}=\begin{bmatrix}\left[\frac{\partial \boldsymbol{f}_1}{\partial \boldsymbol{u}_1}\right]_{\substack{x_1=x_{10}\\u_1=u_{10}}} & & & \boldsymbol{0}\\ & \left[\frac{\partial \boldsymbol{f}_2}{\partial \boldsymbol{u}_2}\right]_{\substack{x_2=x_{20}\\u_2=u_{20}}} & & \\ & & \ddots & \\ \boldsymbol{0} & & & \left[\frac{\partial \boldsymbol{f}_P}{\partial \boldsymbol{u}_P}\right]_{\substack{x_P=x_{P0}\\u_P=u_{P0}}}\end{bmatrix} \tag{23}$$

$$\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{X}}\right]_{\substack{X=X_0\\U=U_0}}=\begin{bmatrix}\left[\frac{\partial \boldsymbol{g}_1}{\partial \boldsymbol{x}_1}\right]_{\substack{x_1=x_{10}\\u_1=u_{10}}} & & & \boldsymbol{0}\\ & \left[\frac{\partial \boldsymbol{g}_2}{\partial \boldsymbol{x}_2}\right]_{\substack{x_2=x_{20}\\u_2=u_{20}}} & & \\ & & \ddots & \\ \boldsymbol{0} & & & \left[\frac{\partial \boldsymbol{g}_P}{\partial \boldsymbol{x}_P}\right]_{\substack{x_P=x_{P0}\\u_P=u_{P0}}}\end{bmatrix} \tag{24}$$

$$\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}=\begin{bmatrix}\left[\frac{\partial \boldsymbol{g}_1}{\partial \boldsymbol{u}_1}\right]_{\substack{x_1=x_{10}\\u_1=u_{10}}} & & & \boldsymbol{0}\\ & \left[\frac{\partial \boldsymbol{g}_2}{\partial \boldsymbol{u}_2}\right]_{\substack{x_2=x_{20}\\u_2=u_{20}}} & & \\ & & \ddots & \\ \boldsymbol{0} & & & \left[\frac{\partial \boldsymbol{g}_P}{\partial \boldsymbol{u}_P}\right]_{\substack{x_P=x_{P0}\\u_P=u_{P0}}}\end{bmatrix}. \tag{25}$$

也就是说，A-I-4 中 $(8a)'$，$(8b)'$ 式的 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$, 在这里是

$$\left[\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{X}}\right]_{\substack{X=X_0\\U=U_0}},\ \left[\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}},\ \left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{X}}\right]_{\substack{X=X_0\\U=U_0}},\ \left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}.$$

此外，因为这里所讨论的线性大系统是，把原来的非线性大系统的各子系统用其线性化子系统替换后得到的，所以各方框以及整个系统相互之间的输入输出连接关系也可以用和(5)，(6)式相同的形式表示．亦即，若令 $\boldsymbol{V}_d$ 和 $\boldsymbol{Z}_d$ 为该线性大系统的输入和输出，则

$$\boldsymbol{U}_d=\boldsymbol{F}\boldsymbol{Y}_d+\boldsymbol{G}\boldsymbol{V}_d \tag{26}$$

$$\boldsymbol{Z}_d=\boldsymbol{J}\boldsymbol{Y}_d+\boldsymbol{K}\boldsymbol{V}_d. \tag{27}$$

这样，线性大系统可以用(20)，(21)，(26)，(27)式完全表示．

可是，根据 A-I-4 中定理 3，若该线性系统能用以 $\boldsymbol{X}_d$ 为状态变量的标准形完全表示，则

$$\boldsymbol{I}-\boldsymbol{F}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}\ \text{是正则的.}$$

由此是否能够说对于原来的非线性大系统也是如此呢？我们再来看一下(13)式

$$\boldsymbol{U}-\boldsymbol{F}\boldsymbol{g}(\boldsymbol{X},\ \boldsymbol{U})=\boldsymbol{G}\boldsymbol{V}. \tag{13}$$

因为将(17)式代入(14)式后，得

$$\boldsymbol{U}_0-\boldsymbol{F}\boldsymbol{g}(\boldsymbol{X}_0,\ \boldsymbol{U}_0)=\boldsymbol{G}\boldsymbol{V}_0, \tag{28}$$

则(13)式具有解 $\boldsymbol{U}_0$, $\boldsymbol{X}_0$, $\boldsymbol{V}_0$．而且，因为 $\boldsymbol{I}-\boldsymbol{F}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}$ 是正则的，根据隐函数定

理[58]，(13)式在 $\boldsymbol{U}_0$, $\boldsymbol{X}_0$, $\boldsymbol{V}_0$ 附近可以解成

$$\boldsymbol{U}=\boldsymbol{h}(\boldsymbol{X},\ \boldsymbol{V}),\ \boldsymbol{h}\in C^1 \tag{9}$$

的形式．其结果，如前所述，在 $\boldsymbol{X}_0$, $\boldsymbol{V}_0$, $\boldsymbol{Z}_0$, $\boldsymbol{U}_0$, $\boldsymbol{Y}_0$ 的附近，$\boldsymbol{X}$, $\boldsymbol{Y}$, $\boldsymbol{Z}$ 也可以表示成

$$\dot{\boldsymbol{X}}=\boldsymbol{f}(\boldsymbol{X},\ \boldsymbol{h}(\boldsymbol{X},\ \boldsymbol{V})) \tag{10}$$

$$\boldsymbol{Y}=\boldsymbol{g}(\boldsymbol{X},\ \boldsymbol{h}(\boldsymbol{X},\ \boldsymbol{V})) \tag{11}$$

$$\boldsymbol{Z}=\boldsymbol{J}\boldsymbol{g}(\boldsymbol{X},\ \boldsymbol{h}(\boldsymbol{X},\ \boldsymbol{V}))+\boldsymbol{K}\boldsymbol{V} \tag{12}$$

的形式．亦即，若由线性化后的子系统组成的大系统可以以 $\boldsymbol{X}_d$ 为状态变量，用标准形完全表示，则原来的非线性大系统可以在 $\boldsymbol{X}_0$, $\boldsymbol{V}_0$, $\boldsymbol{Z}_0$, $\boldsymbol{U}_0$, $\boldsymbol{Y}_0$ 的附近用正规形完全表示．因此必须注意，在这个平衡解附近具有唯一解．

那么，利用 A-I-4 中(11a)，(11b)式，表示由该线性化子系统组成的大系统的标准形变成

$$\begin{aligned}\dot{\boldsymbol{X}}_d=&\left\{\left[\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{X}}\right]_{\substack{\boldsymbol{X}=\boldsymbol{X}_0\\ \boldsymbol{U}=\boldsymbol{U}_0}}+\left[\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{U}}\right]_{\substack{\boldsymbol{X}=\boldsymbol{X}_0\\ \boldsymbol{U}=\boldsymbol{U}_0}}\left(\boldsymbol{I}-\boldsymbol{F}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{\boldsymbol{X}=\boldsymbol{X}_0\\ \boldsymbol{U}=\boldsymbol{U}_0}}\right)^{-1}\boldsymbol{F}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{X}}\right]_{\substack{\boldsymbol{X}=\boldsymbol{X}_0\\ \boldsymbol{U}=\boldsymbol{U}_0}}\right\}\boldsymbol{X}_d\\ &+\left[\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{U}}\right]_{\substack{\boldsymbol{X}=\boldsymbol{X}_0\\ \boldsymbol{U}=\boldsymbol{U}_0}}\left(\boldsymbol{I}-\boldsymbol{F}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{\boldsymbol{X}=\boldsymbol{X}_0\\ \boldsymbol{U}=\boldsymbol{U}_0}}\right)^{-1}\boldsymbol{G}\boldsymbol{V}_d\end{aligned} \tag{29}$$

$$\begin{aligned}\boldsymbol{Z}_d=&\boldsymbol{J}\left(\boldsymbol{I}-\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{\boldsymbol{X}=\boldsymbol{X}_0\\ \boldsymbol{U}=\boldsymbol{U}_0}}\boldsymbol{F}\right)^{-1}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{X}}\right]_{\substack{\boldsymbol{X}=\boldsymbol{X}_0\\ \boldsymbol{U}=\boldsymbol{U}_0}}\boldsymbol{X}_d\\ &+\left\{\boldsymbol{J}\left(\boldsymbol{I}-\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{\boldsymbol{X}=\boldsymbol{X}_0\\ \boldsymbol{U}=\boldsymbol{U}_0}}\boldsymbol{F}\right)^{-1}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{\boldsymbol{X}=\boldsymbol{X}_0\\ \boldsymbol{U}=\boldsymbol{U}_0}}\boldsymbol{G}+\boldsymbol{K}\right\}\boldsymbol{V}_d\end{aligned} \tag{30}$$

这个标准形是否与表示原来的非线性系统(在 $\boldsymbol{X}_0$, $\boldsymbol{V}_0$, $\boldsymbol{Z}_0$, $\boldsymbol{U}_0$, $\boldsymbol{Y}_0$ 的附近)的正规形的线性化表示式一致呢？为了研究这个问题，试将(10)，(12)式在 $\boldsymbol{X}_0$, $\boldsymbol{V}_0$, $\boldsymbol{Z}_0$ 的附近线性化．在这里，设 $\tilde{\boldsymbol{X}}_d$, $\tilde{\boldsymbol{V}}_d$, $\tilde{\boldsymbol{Z}}_d$ 分别为 $\boldsymbol{X}$, $\boldsymbol{V}$, $\boldsymbol{Z}$ 相对于 $\boldsymbol{X}_0$, $\boldsymbol{V}_0$, $\boldsymbol{Z}_0$ 的变化量，则

$$\dot{\tilde{\boldsymbol{X}}}_d=\left(\left[\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{X}}\right]_{\substack{\boldsymbol{X}=\boldsymbol{X}_0\\ \boldsymbol{U}=\boldsymbol{U}_0}}+\left[\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{U}}\right]_{\substack{\boldsymbol{X}=\boldsymbol{X}_0\\ \boldsymbol{U}=\boldsymbol{U}_0}}\left[\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{X}}\right]_{\substack{\boldsymbol{X}=\boldsymbol{X}_0\\ \boldsymbol{V}=\boldsymbol{V}_0}}\right)\tilde{\boldsymbol{X}}_d+\left[\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{U}}\right]_{\substack{\boldsymbol{X}=\boldsymbol{X}_0\\ \boldsymbol{U}=\boldsymbol{U}_0}}\left[\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{V}}\right]_{\substack{\boldsymbol{X}=\boldsymbol{X}_0\\ \boldsymbol{V}=\boldsymbol{V}_0}}\tilde{\boldsymbol{V}}_d \tag{31}$$

$$\begin{aligned}\tilde{\boldsymbol{Z}}_d=&\boldsymbol{J}\left(\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{X}}\right]_{\substack{\boldsymbol{X}=\boldsymbol{X}_0\\ \boldsymbol{U}=\boldsymbol{U}_0}}+\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{\boldsymbol{X}=\boldsymbol{X}_0\\ \boldsymbol{U}=\boldsymbol{U}_0}}\left[\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{X}}\right]_{\substack{\boldsymbol{X}=\boldsymbol{X}_0\\ \boldsymbol{V}=\boldsymbol{V}_0}}\right)\tilde{\boldsymbol{X}}_d\\ &+\left(\boldsymbol{J}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{\boldsymbol{X}=\boldsymbol{X}_0\\ \boldsymbol{U}=\boldsymbol{U}_0}}\left[\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{V}}\right]_{\substack{\boldsymbol{X}=\boldsymbol{X}_0\\ \boldsymbol{V}=\boldsymbol{V}_0}}+\boldsymbol{K}\right)\tilde{\boldsymbol{V}}_d\end{aligned} \tag{32}$$

在上面两式右边的系数矩阵中，只知道 $\left[\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{X}}\right]_{\substack{\boldsymbol{X}=\boldsymbol{X}_0\\ \boldsymbol{V}=\boldsymbol{V}_0}}$ 和 $\left[\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{V}}\right]_{\substack{\boldsymbol{X}=\boldsymbol{X}_0\\ \boldsymbol{V}=\boldsymbol{V}_0}}$ 存在，是 $\boldsymbol{h}(\circ,\ \circ)$ 的属于 C^1 的函数，而不知道其具体形式，所以不能直接计算．但是，若注意到(9)式和(13)式在 $\boldsymbol{X}_0$, $\boldsymbol{V}_0$, $\boldsymbol{U}_0$ 的附近是等价的，可计算如下．

将(9)式代入(13)式，得

$$\boldsymbol{h}(\boldsymbol{X},\ \boldsymbol{V})-\boldsymbol{F}\boldsymbol{g}(\boldsymbol{X},\ \boldsymbol{h}(\boldsymbol{X},\ \boldsymbol{V}))-\boldsymbol{G}\boldsymbol{V}=\boldsymbol{0} \tag{33}$$

因(9)式和(13)式是等价的，则该式不是方程式，而是恒等式，即对于任意 $\boldsymbol{X}$, $\boldsymbol{V}$(在 $\boldsymbol{X}_0$ 和 $\boldsymbol{V}_0$ 的附近)都成立．因此，将(33)式在 $\boldsymbol{X}=\boldsymbol{X}_0$, $\boldsymbol{V}=\boldsymbol{V}_0$ 处全微分，得

$$\left\{\left[\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{X}}\right]_{\substack{X=X_0\\V=V_0}}-\boldsymbol{F}\left(\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{X}}\right]_{\substack{X=X_0\\U=U_0}}+\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}\left[\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{X}}\right]_{\substack{X=X_0\\V=V_0}}\right)\right\}d\boldsymbol{X}$$

$$+\left\{\left[\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{V}}\right]_{\substack{X=X_0\\V=V_0}}-\boldsymbol{F}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}\left[\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{V}}\right]_{\substack{X=X_0\\V=V_0}}-\boldsymbol{G}\right\}d\boldsymbol{V}=\boldsymbol{0} \tag{34}$$

上式中 $d\boldsymbol{X}$ 和 $d\boldsymbol{V}$ 的系数矩阵必须为零矩阵．即

$$\left[\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{X}}\right]_{\substack{X=X_0\\V=V_0}}-\boldsymbol{F}\left(\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{X}}\right]_{\substack{X=X_0\\U=U_0}}+\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}\left[\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{X}}\right]_{\substack{X=X_0\\V=V_0}}\right)=\boldsymbol{0}$$

$$\left[\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{V}}\right]_{\substack{X=X_0\\V=V_0}}-\boldsymbol{F}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}\left[\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{V}}\right]_{\substack{X=X_0\\V=V_0}}-\boldsymbol{G}=\boldsymbol{0}$$

将上列两式整理后得

$$\left(\boldsymbol{I}-\boldsymbol{F}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}\right)\left[\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{V}}\right]_{\substack{X=X_0\\V=V_0}}=\boldsymbol{F}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{X}}\right]_{\substack{X=X_0\\U=U_0}}$$

$$\left(\boldsymbol{I}-\boldsymbol{F}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}\right)\left[\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{V}}\right]_{\substack{X=X_0\\V=V_0}}=\boldsymbol{G}.$$

因 $\boldsymbol{I}-\boldsymbol{F}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}$ [1)]是正则矩阵，则上述 $\left[\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{X}}\right]_{\substack{X=X_0\\V=V_0}}$， $\left[\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{V}}\right]_{\substack{X=X_0\\V=V_0}}$ 可按下列两式计算

$$\left[\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{X}}\right]_{\substack{X=X_0\\V=V_0}}=\left(\boldsymbol{I}-\boldsymbol{F}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}\right)^{-1}\boldsymbol{F}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{X}}\right]_{\substack{X=X_0\\U=U_0}} \tag{35}$$

$$\left[\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{V}}\right]_{\substack{X=X_0\\V=V_0}}=\left(\boldsymbol{I}-\boldsymbol{F}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}\right)^{-1}\boldsymbol{G} \tag{36}$$

根据以上，表示非线性大系统正规形(10)，(12)式的线性化表达式变成

$$\dot{\tilde{\boldsymbol{X}}}_d=\left\{\left[\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{X}}\right]_{\substack{X=X_0\\U=U_0}}+\left[\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}\left(\boldsymbol{I}-\boldsymbol{F}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}\right)^{-1}\boldsymbol{F}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{X}}\right]_{\substack{X=X_0\\U=U_0}}\right\}\tilde{\boldsymbol{X}}_d$$

$$+\left[\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}\left(\boldsymbol{I}-\boldsymbol{F}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}\right)^{-1}\boldsymbol{G}\tilde{\boldsymbol{V}}_d \tag{37}$$

$$\tilde{\boldsymbol{Z}}_d=\boldsymbol{J}\left\{\boldsymbol{I}+\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}\left(\boldsymbol{I}-\boldsymbol{F}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}\right)^{-1}\boldsymbol{F}\right\}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{X}}\right]_{\substack{X=X_0\\U=U_0}}\tilde{\boldsymbol{X}}_d$$

$$+\left\{\boldsymbol{J}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}\left(\boldsymbol{I}-\boldsymbol{F}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}\right)^{-1}\boldsymbol{G}+\boldsymbol{K}\right\}\tilde{\boldsymbol{V}}_d \tag{38}$$

将上列两式与表示由线性化子系统组成的大系统的标准形(29)，(30)比较，可见(37)式与(29)式完全相同．而且，因为一般等式

$$\boldsymbol{I}+\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}\left(\boldsymbol{I}-\boldsymbol{F}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}\right)^{-1}\boldsymbol{F}=\left(\boldsymbol{I}-\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}\boldsymbol{F}\right)^{-1}$$

$$\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}\left(\boldsymbol{I}-\boldsymbol{F}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}\right)^{-1}=\left(\boldsymbol{I}-\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}\boldsymbol{F}\right)^{-1}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\U=U_0}}$$

1) 原文误为 $\boldsymbol{I}-\boldsymbol{F}\left[\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{U}}\right]_{\substack{X=X_0\\V=V_0}}$．——译者注

成立，则(38)式也与(30)式完全相同．亦即，若由线性化子系统组成的大系统可以以各子系统状态变量的直积为状态变量，用标准形完全表示，则其标准形就是表示原来非线性大系统正规形的线性化表示式．

以上讨论，是假定有满足(3)～(6)式的常数向量解 $\boldsymbol{X}_0$, $\boldsymbol{U}_0$, $\boldsymbol{Z}_0$, $\boldsymbol{U}_0$, $\boldsymbol{Y}_0$ 存在，但是在满足(3)～(6)式的任意解(不一定是常数向量)附近，这些讨论也同样成立，即可以得到下面定理．

[定理 4]

设由非线性子系统组成的大系统具有一个解．将各子系统换成它在该解附近的线性化子系统，现来考察所得到的线性大系统．若该线性大系统可以以各子系统状态变量的直积为状态变量，用标准形完全表示，则由原来的非线性子系统组成的大系统，在上述解附近具有唯一解．而且，该线性化表示式(在上述解附近得到的)和表示由线性化子系统组成的线性大系统的标准形一致．

由线性化子系统组成的大系统，在不能用以各子系统状态变量的直积为状态变量的标准形表示的情况下，该定理不能说明任何问题．因此，它还不能说是完全回答了 $Q1$ 的问题．但是对于实际问题，即使只是上述定理，在将非线性大系统的各子系统线性化后组合成的整个系统作为线性系统处理的许多情况下，也可以认为是重要的理论依据．

那末，现用定理 4 来分析一下用方框图表示的一个系统．

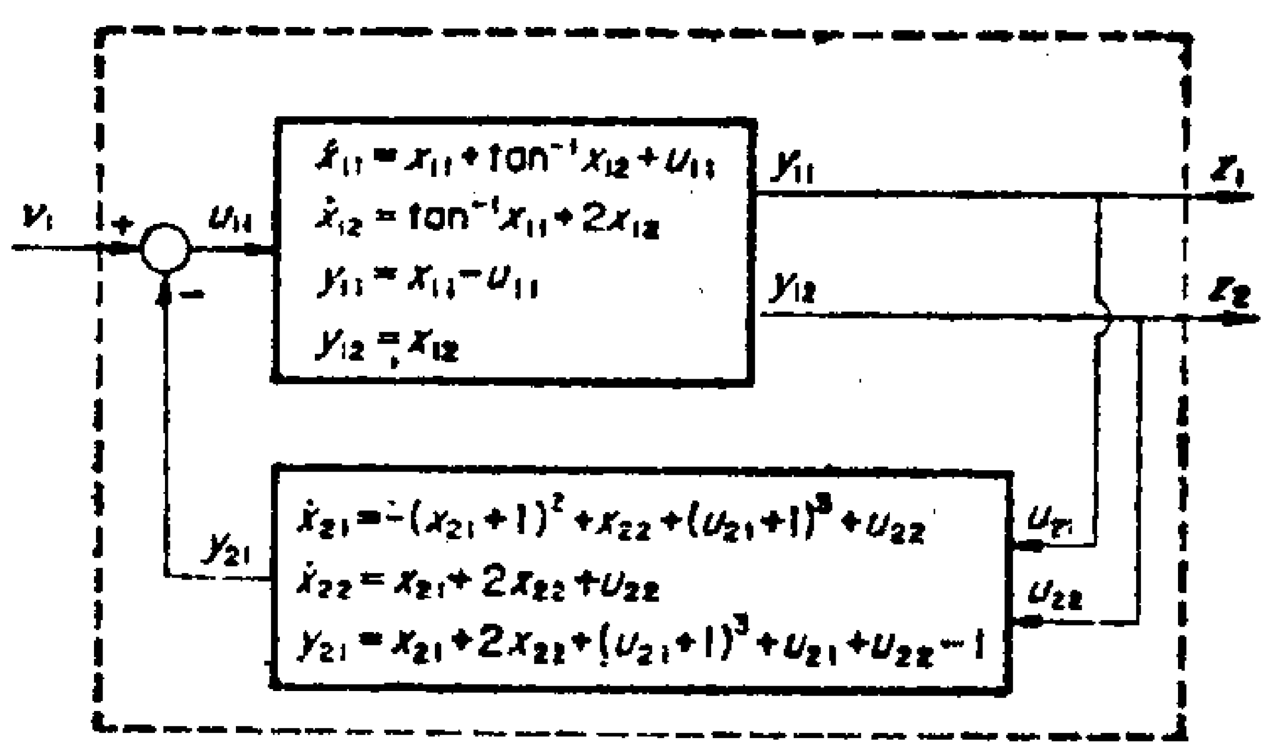


图 1　由非线性子系统组合成的反馈系统

(例 1)　现来看一下图 1 所示系统，其各子系统可以用正规形表示．显然，该系统具有下列常数解．

$$\begin{bmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad V_1 = 0, \quad \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} u_{11} \\ u_{21} \\ u_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \qquad \begin{bmatrix} y_{11} \\ y_{12} \\ y_{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{39}$$

在该解附近将各方框线性化，各方框用其线性化系统替换．对应于 A-I-4 中(8a)′，(8b)′，(5)′，(6)′式．将由该替换得到的线性系统(图 2)表示成

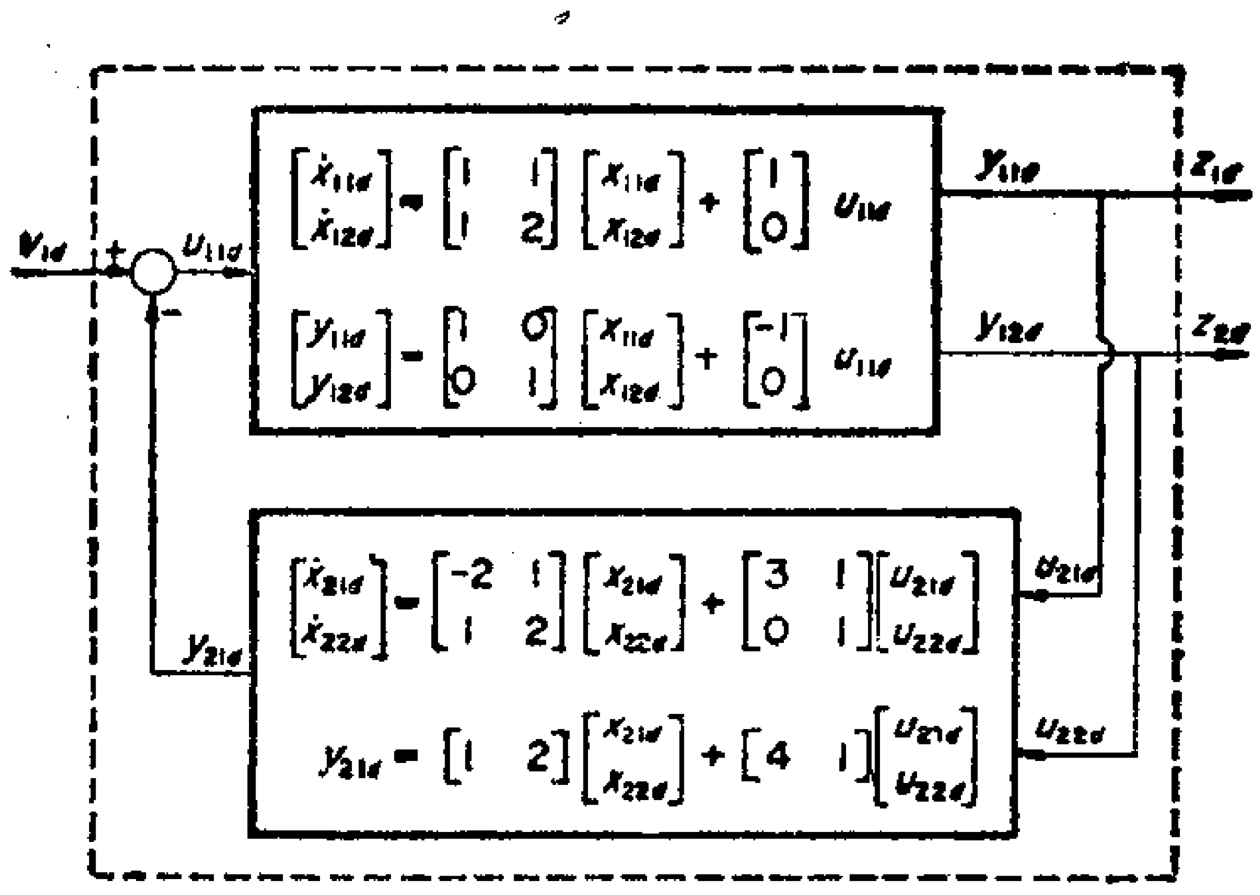


图 2　将图 1 中各子系统线性化得到的反馈系统

$$\begin{bmatrix}\dot{x}_{11d}\\ \dot{x}_{12d}\\ \hdashline \dot{x}_{21d}\\ \dot{x}_{22d}\end{bmatrix}=\underset{\boldsymbol{A}}{\left[\begin{array}{cc:cc}1&1&0&0\\1&2&0&0\\ \hdashline 0&0&-2&1\\0&0&1&2\end{array}\right]}\begin{bmatrix}x_{11d}\\x_{12d}\\ \hdashline x_{21d}\\x_{22d}\end{bmatrix}+\underset{\boldsymbol{B}}{\left[\begin{array}{c:cc}1&0&0\\0&0&0\\ \hdashline 0&3&1\\0&0&1\end{array}\right]}\begin{bmatrix}u_{11d}\\ \hdashline u_{21d}\\u_{22d}\end{bmatrix}\tag{40}$$

$$\begin{bmatrix}y_{11d}\\y_{12d}\\ \hdashline y_{21d}\end{bmatrix}=\underset{\boldsymbol{C}}{\left[\begin{array}{cc:cc}1&0&0&0\\0&1&0&0\\ \hdashline 0&0&1&2\end{array}\right]}\begin{bmatrix}x_{11d}\\x_{12d}\\ \hdashline x_{21d}\\x_{22d}\end{bmatrix}+\underset{\boldsymbol{D}}{\left[\begin{array}{c:cc}-1&0&0\\0&0&0\\ \hdashline 0&4&1\end{array}\right]}\begin{bmatrix}u_{11d}\\ \hdashline u_{21d}\\u_{22d}\end{bmatrix}\tag{41}$$

$$\begin{bmatrix}u_{11d}\\ \hdashline u_{21d}\\u_{22d}\end{bmatrix}=\underset{\boldsymbol{F}}{\left[\begin{array}{cc:c}0&0&-1\\ \hdashline 1&0&0\\0&1&0\end{array}\right]}\begin{bmatrix}y_{11d}\\y_{12d}\\ \hdashline y_{21d}\end{bmatrix}+\underset{\boldsymbol{G}}{\begin{bmatrix}1\\ \hdashline 0\\0\end{bmatrix}}V_{1d}\tag{42}$$

$$\begin{bmatrix}Z_{1d}\\Z_{2d}\end{bmatrix}=\underset{\boldsymbol{J}}{\left[\begin{array}{cc:c}1&0&0\\0&1&0\end{array}\right]}\begin{bmatrix}y_{11d}\\y_{12d}\\ \hdashline y_{21d}\end{bmatrix}+\underset{\boldsymbol{K}}{\begin{bmatrix}0\\0\end{bmatrix}}V_{1d}\tag{43}$$

计算 $\det[\boldsymbol{I}-\boldsymbol{FD}]$，得

$$\det\begin{bmatrix}1&4&1\\1&1&0\\0&0&1\end{bmatrix}=-3\neq0$$

根据 A-I-4 中定理 3，则可保证该线性系统以 $[x_{11d},\ x_{12d},\ x_{21d},\ x_{22d}]^T$ 为状态变量能用标准形完全表示. 因此根据定理 4，图 1 所示非线性系统在 (39) 式所示解的附近具有唯一解，可以用 (7)，(8) 式形式正规形完全表示. 下面，我们来求一下它的正规形. 首先，将与 (13) 式对应的

$$\begin{bmatrix}u_{11}\\u_{21}\\u_{22}\end{bmatrix}-\begin{bmatrix}0&0&-1\\1&0&0\\0&1&0\end{bmatrix}\begin{bmatrix}x_{11}-u_{11}\\x_{12}\\x_{21}+2x_{22}+(u_{21}+1)^3+u_{21}+u_{22}-1\end{bmatrix}=\begin{bmatrix}1\\0\\0\end{bmatrix}V_1\tag{44}$$

展开，得

$$u_{11}=-x_{21}-2x_{22}-(u_{21}+1)^3-u_{21}-u_{22}+1+V_1 \tag{45a}$$

$$u_{21}=x_{11}-u_{11} \tag{45b}$$

$$u_{22}=x_{12}. \tag{45c}$$

u_{22} 已经解成(45c)的形式. 将(45b), (45c)式代入(45a)式中的 u_{21}, u_{22} 项($(u_{21}+1)^3$ 项原样不动)，得

$$(u_{21}+1)^3=-x_{11}-x_{12}-x_{21}-2x_{22}+V_1+1$$

对于 u_{21} 可解成

$$u_{21}=-(x_{11}+x_{12}+x_{21}+2x_{22}-V_1-1)^{1/3}-1 \tag{46}$$

因此，由(45b)式得

$$u_{11}=x_{11}+(x_{11}+x_{12}+x_{21}+2x_{22}-V_1-1)^{1/3}+1 \tag{47}$$

将这些表示式代入表示各子系统的正规形，得

$$\left.\begin{aligned}\dot{x}_{11}&=2x_{11}+\tan^{-1}x_{12}+(x_{11}+x_{12}+x_{21}+2x_{22}-V_1-1)^{1/3}+1\\ \dot{x}_{12}&=\tan^{-1}x_{11}+2x_{12}\\ \dot{x}_{21}&=-x_{11}-x_{21}^2-3x_{21}-x_{22}+V_1\\ \dot{x}_{22}&=x_{12}+x_{21}+2x_{22}\end{aligned}\right\} \tag{48}$$

$$\left.\begin{aligned}Z_1(=y_{11})&=-(x_{11}+x_{12}+x_{21}+2x_{22}-V_1-1)^{1/3}-1\\ Z_2(=y_{12})&=x_{12}.\end{aligned}\right\} \tag{49}$$

因为上列各式的右边在 $V_1=0$, $[x_{11},\ x_{12},\ x_{21},\ x_{22}]^T=[0\ 0\ 0\ 0]^T$ 的附近确实连续可微，所以就具体地证实了在这个附近整个系统可以用(7)，(8)式形式的正规形表示[1]. 而且，(48)，(49)在的 $V_1=0$, $[x_{11},\ x_{12},\ x_{21},\ x_{22}]^T=[0\ 0\ 0\ 0]^T$, $[Z_1,\ Z_2]^T=[0\ 0]^T$ 近旁的，线性化表示式变成

$$\begin{bmatrix}\tilde{x}_{11d}\\ \tilde{x}_{12d}\\ \tilde{x}_{21d}\\ \tilde{x}_{22d}\end{bmatrix}=\frac{1}{3}\begin{bmatrix}7&4&1&2\\ 3&6&0&0\\ -3&0&-9&-3\\ 0&3&3&6\end{bmatrix}\begin{bmatrix}\tilde{x}_{11d}\\ \tilde{x}_{12d}\\ \tilde{x}_{21d}\\ \tilde{x}_{22d}\end{bmatrix}+\frac{1}{3}\begin{bmatrix}-1\\ 0\\ 3\\ 0\end{bmatrix}\tilde{V}_{1d} \tag{50}$$

$$\begin{bmatrix}\tilde{Z}_{1d}\\ \tilde{Z}_{2d}\end{bmatrix}=\frac{1}{3}\begin{bmatrix}-1&-1&-1&-2\\ 0&3&0&0\end{bmatrix}\begin{bmatrix}\tilde{x}_{11d}\\ \tilde{x}_{12d}\\ \tilde{x}_{21d}\\ \tilde{x}_{22d}\end{bmatrix}+\frac{1}{3}\begin{bmatrix}1\\ 0\end{bmatrix}\tilde{V}_{1d} \tag{51}$$

它与利用 A-I-4 中(11a)′, (11b)′ 式，由(40)～(43)式计算出的表示图 2 线性系统的标准形一致.

1) (48)、(49)式不是到处都连续可微(例如，试考虑一下在 $V_1=-1$, $[x_{11},\ x_{12},\ x_{21},\ x_{22}]^T=[0\ 0\ 0\ 0]^T$ 处的微分可能性)，因此不能说整个系统能够大范围地用(7)、(8)式形式正规形表示.

B-I-9 克兰姆行列式

下一章将要讨论一般的(时变)线性系统. 那时, 常常会出现以实轴区间$[t_0, t_1]$上定义的连续实函数$f_{ij}(\circ)$为元素的矩阵

$$F(t) \triangleq [f_{ij}(t)] = m \times n \tag{1}$$

如本书开始B-0-2中所述,上述连续实函数是实数体R上的向量空间. 该向量空间我们用$C[t_0, t_1]$表示,则$F(\cdot)$的列(行)可以看成是向量空间$C[t_0, t_1]^m(C[t_0, t_1]^n)$的元. 在下一章以后的分析中象这样考虑时,往往会出现下面的问题, $F(\cdot)$的列(行)是线性独立的,还是线性相关.

必须反复强调,若属于$C[t_0, t_1]^m$的n个向量$f_j(\cdot)\ (j=1, \cdots, n)$线性相关, 则对于某些不全为0的实数$\alpha_i(i=1, \cdots, n)$, 下式成立

$$\alpha_1 f_1(t) + \cdots + \alpha_n f_n(t) = 0, \quad \forall t \in [t_0, t_1] \tag{2}$$

这里必须注意下面两点,第一,对于所有的$t \in [t_0, t_1]$, 上式必须恒等于$\mathbf{0}$. 其次, α_i不是t的函数,必须是常数.

关于判断所给一组向量是否线性独立的方法, 在B-I-3中已经讲了一点. 在向量元所属的集合和数所属的集合在同一个体F的情况下, 向量$a_1, \cdots, a_n$是否线性独立, 可以通过以这些向量为列组成的矩阵

$$A \triangleq [a_1, \cdots, a_n] \tag{3}$$

的秩判定. 但是若注意一下上面讲到的例子, 在向量元所属集合和数所属的体不同的情况下,这个判别法不能使用. 上述$C[t_0, t]^m$就相当于这种情况. 因此, 为了知道属于$C[t_0, t]^m$的一组向量是否线性独立, 必须使用其它判别方法. 下面介绍的利用克兰姆行列式的判别方法就是其中的一个.

[定义] 令$f_k(\cdot)\ (k=1, \cdots, n)$为$C[t_0, t_1]^m$的向量时, 以

$$g_{ij} \triangleq \int_{t_0}^{t_1} f_i^T(t) f_j(t)\, dt \quad i, j=1, \cdots, n \tag{4}$$

为$i-j$元素的矩阵

$$G \triangleq [g_{ij}] \in R^{n\times n} \tag{5}$$

称为克兰姆矩阵(Gram matrix), $\det G$称为克兰姆行列式(Gramian).

(问题1) 试证明克兰姆矩阵是对称矩阵,是准正定矩阵.

(问题2) 试证明与(1)式矩阵$F(t)$的n个列向量对应的克兰姆矩阵可以用下式给出:

$$G = \int_{t_0}^{t_1} F^T(t) F(t)\, dt \tag{6}$$

利用该定义,则下面定理成立.

[定理1] $f_k(\cdot)\ (k=1, \cdots, n)$线性独立的充分和必要条件是$\det G \neq 0$, 即克兰姆矩阵是正则的.

[证明] 在(2)式两边左乘以$f_i^T(t)$后,对t积分,得

$$\int_{t_0}^{t_1} \boldsymbol{f}_i^T(t) \sum_{k=1}^{n} \alpha_k \boldsymbol{f}_k(t) dt = 0 \quad i=1, 2, \cdots, n \tag{7}$$

将(7)式左边改变形式后，得

$$\int_{t_0}^{t_1} \boldsymbol{f}_i^T(t) \sum_{k=1}^{n} \alpha_k \boldsymbol{f}_k(t) dt = \sum_{k=1}^{n} \alpha_k \int_{t_0}^{t_1} \boldsymbol{f}_i^T(t) \boldsymbol{f}_k(t) dt = \sum_{k=1}^{n} g_{ik}\alpha_k \quad i=1, 2, \cdots, n \tag{8}$$

将(8)式对所有的 i 全部归纳后，写成

$$\boldsymbol{G\alpha} = \boldsymbol{0} \tag{9}$$

式中 $\boldsymbol{\alpha}$ 是以 α_k 为元素的列向量，$\boldsymbol{\alpha} \triangleq [\alpha_1, \alpha_2, \cdots, \alpha_n]^T$. 若 $\boldsymbol{G}$ 是正则的，则(9)式解仅限于 $\boldsymbol{\alpha} = \boldsymbol{0}$(B-I-4). 因此，$\boldsymbol{f}_k(t)$ $(k=1, 2, \cdots, n)$是线性独立的.

若 $\boldsymbol{G}$ 不是正则矩阵，则有满足(9)式的 $\boldsymbol{\alpha} \neq \boldsymbol{0}$ 存在. 因此，对于不全为 0 的 $\alpha_k(k=1, 2, \cdots, n)$，(7)式成立. 将(7)式的第 1 式两边乘以 α_1，第 2 式两边乘以 α_2，然后加起来，得

$$\sum_{i=1}^{n} \alpha_i \int_{t_0}^{t_1} \boldsymbol{f}_i^T(t) \sum_{k=1}^{n} \alpha_k \boldsymbol{f}_k(t) dt = 0 \tag{10}$$

上式整理后，得

$$\sum_{i=1}^{n} \alpha_i \int_{t_0}^{t_1} \boldsymbol{f}_i^T(t) \sum_{k=1}^{n} \alpha_k \boldsymbol{f}_k(t) dt = \int_{t_0}^{t_1} \left[\sum_{i=1}^{n} \alpha_i \boldsymbol{f}_i^T(t) \right] \left[\sum_{k=1}^{n} \alpha_k \boldsymbol{f}_k(t) \right] dt \tag{11}$$

现设 $\boldsymbol{f}(t) \triangleq \sum_{i=1}^{n} \alpha_i \boldsymbol{f}_i(t)$，将(11)式代入(10)式，得

$$\int_{t_0}^{t_1} \boldsymbol{f}^T(t) \boldsymbol{f}(t) dt = 0. \tag{12}$$

因为 $\boldsymbol{f}(\cdot)$ 是连续的，$\boldsymbol{f}^T(\cdot)\boldsymbol{f}(\cdot)$ 对于所有的 t 都是非负的，则只有在

$$\boldsymbol{f}(t) = \sum_{i=1}^{n} \alpha \boldsymbol{f}_i(t) = \boldsymbol{0}, \ \forall t \in [t_0, t_1] \tag{13}$$

的情况下(12)式才成立. 因而 $\boldsymbol{f}_i(\cdot)$是线性相关的.

(注意)

定理 1 可以推广到在区间$[t_0, t_1]$上定义的分段连续的复变函数的场合. 亦即，当设矩阵 $\boldsymbol{F}(t) \triangleq [f_{ij}(t)] = m \times n$ 的各元素 $f_{ij}(\cdot)$ 是在$[t_0, t_1]$上分段连续，具有复数值的函数时，$\boldsymbol{F}(\cdot)$ 的 n 个列向量在$[t_0, t_1]$上线性独立的充分和必要条件是，克兰姆矩阵

$$\boldsymbol{G} = \int_{t_0}^{t_1} \boldsymbol{F}^*(t) \boldsymbol{F}(t) dt$$

为正则矩阵. 式中 $\boldsymbol{F}^*(t)$ 是 $\boldsymbol{F}(t)$ 的共轭转置矩阵.

向量的内积

现在我们规定，对两个向量 $\boldsymbol{f}_i(\cdot)$和 $\boldsymbol{f}_j(\cdot)$进行(4)式右边的计算时，用符号 $\langle \boldsymbol{f}_i, \boldsymbol{f}_j \rangle$ 表示. 在上面定理的证明中我们发现，仅仅用到关于运算$\langle \cdot, \cdot \rangle$的下列四个性质:

(i) $\langle \boldsymbol{f}_i, \boldsymbol{f}_j \rangle = \langle \boldsymbol{f}_j, \boldsymbol{f}_i \rangle$

(ii) 设 λ 为实数. 则 $\langle \lambda \boldsymbol{f}_i, \boldsymbol{f}_j \rangle = \lambda \langle \boldsymbol{f}_i, \boldsymbol{f}_j \rangle$

(iii) $\langle \boldsymbol{f}_k, \boldsymbol{f}_i + \boldsymbol{f}_j \rangle = \langle \boldsymbol{f}_k, \boldsymbol{f}_i \rangle + \langle \boldsymbol{f}_k, \boldsymbol{f}_j \rangle$

(iv) $\langle \boldsymbol{f}_i, \boldsymbol{f}_i \rangle^* = 0 \Leftrightarrow \boldsymbol{f}_i = 0$.

* 原文误印为 $\langle f_i, f_i \rangle \Leftrightarrow f_i = 0$. ——译者注

设 V 是体 $C(R)$ 上的向量空间. 当 V 的两个向量 $\boldsymbol{x}, \boldsymbol{y}$ 按 $\langle\boldsymbol{x}, \boldsymbol{y}\rangle$ 与某个复数(实数)相对应,而且运算 $\langle\cdot, \cdot\rangle$ 满足下列公理

(i) $\langle\boldsymbol{x}, \boldsymbol{y}\rangle=\overline{\langle\boldsymbol{y}, \boldsymbol{x}\rangle}$ ("——"表示共轭复数)

(ii) $\langle\lambda\boldsymbol{x}, \boldsymbol{y}\rangle=\lambda\langle\boldsymbol{x}, \boldsymbol{y}\rangle$

(iii) $\langle\boldsymbol{x}+\boldsymbol{y}, \boldsymbol{z}_k\rangle=\langle\boldsymbol{x}, \boldsymbol{z}_k\rangle+\langle\boldsymbol{y}, \boldsymbol{z}_k\rangle$

(iv) $\langle\boldsymbol{x}, \boldsymbol{x}\rangle\geqslant 0$, 且 $\langle\boldsymbol{x}, \boldsymbol{x}\rangle=0\Leftrightarrow\boldsymbol{x}=\mathbf{0}$

时, $\langle\cdot, \cdot\rangle$ 称为内积(inner product), 可以定义内积的空间称为内积空间(inner product space).

由上面的证明过程可以看到, 定理 1 不仅适用于 $C[t_0, t_1]^m$, 而且可以推广到一般的内积空间.

(问题 3) 对于体 C 上的向量空间,试证明克兰姆矩阵 $\boldsymbol{G}$ 是埃尔米特阵.

(问题 4) 对于向量空间 R^m 中的向量 $\boldsymbol{a}_i$, 若定义

$$\langle\boldsymbol{a}_i, \boldsymbol{a}_j\rangle\triangleq\boldsymbol{a}_i^T\boldsymbol{a}_j,$$

则满足上述性质(i)～(iv). 因此, $\boldsymbol{a}_i^T\boldsymbol{a}_j$ 是 R^m 中的内积,试证明之.

(问题 5) 利用问题 4 的结果,作对应于 $\boldsymbol{a}_i\in R^m(i=1, \cdots, n)$ 的克兰姆矩阵 $\boldsymbol{G}$,试证明它和(3)式的矩阵 $\boldsymbol{A}$ 之间存在下列关系

$$\boldsymbol{G}=\boldsymbol{A}^T\boldsymbol{A}.$$

并由此证明,按 $\boldsymbol{A}$ 的秩判断线性独立性和克兰姆矩阵方法是等价的.

这里,关于内积和 B-I-5 中讲过的范数之间的关系作一些说明.

现在假定, 在某向量空间 $V(F)$ (F 是 C 或 R) 中, 可以定义满足公理(i)～(iv)的内积. 利用该内积, 对任意 $\boldsymbol{x}\in V$ 考察纯量 $N(\boldsymbol{x})\triangleq\langle\boldsymbol{x}, \boldsymbol{x}\rangle^{1/2}$ 时,可以证明,该量完全满足 B-I-5 中讲过的范数的定义,即

(i) $N(\boldsymbol{x})>0$, ${}^\forall\boldsymbol{x}\neq\theta$, $N(\theta)=0$

(ii) $N(\alpha\boldsymbol{x})=|\alpha|N(\boldsymbol{x})$, ${}^\forall\alpha\in F$, ${}^\forall\boldsymbol{x}\in V$

(iii) $N(\boldsymbol{x}+\boldsymbol{y})\leqslant N(\boldsymbol{x})+N(\boldsymbol{y})$, ${}^\forall\boldsymbol{x}, \boldsymbol{y}\in V$

成立.

换句话说,下面结论成立.

[定理 2] 在可以定义内积的向量空间中,函数 $\|\boldsymbol{x}\|=\langle\boldsymbol{x}, \boldsymbol{x}\rangle^{1/2}$ 满足范数条件,即内积空间常可按 $\|\boldsymbol{x}\|=\langle\boldsymbol{x}, \boldsymbol{x}\rangle^{1/2}$ 形成赋范空间.

(证明) 只要证明 $\|\boldsymbol{x}\|=\langle\boldsymbol{x}, \boldsymbol{x}\rangle^{1/2}$ 满足上述范数条件(i), (ii), (iii)即可. 因为(i), (ii)显然,现只证明(iii). 首先

$$\begin{aligned}\|\boldsymbol{x}+\boldsymbol{y}\|^2&=\langle\boldsymbol{x}+\boldsymbol{y}, \boldsymbol{x}+\boldsymbol{y}\rangle=\langle\boldsymbol{x}, \boldsymbol{x}\rangle+\langle\boldsymbol{x}, \boldsymbol{y}\rangle+\langle\boldsymbol{y}, \boldsymbol{x}\rangle\\&\quad+\langle\boldsymbol{y}, \boldsymbol{y}\rangle\leqslant\|\boldsymbol{x}\|^2+2|\langle\boldsymbol{x}, \boldsymbol{y}\rangle|+\|\boldsymbol{y}\|^2\end{aligned}\tag{13}$$

成立. 而且,若利用许瓦尔兹不等式[1](Schwarz inequality)

$$|\langle\boldsymbol{x}, \boldsymbol{y}\rangle|\leqslant\|\boldsymbol{x}\|\cdot\|\boldsymbol{y}\|.\tag{14}$$

(13)式又变成

$$\|\boldsymbol{x}+\boldsymbol{y}\|^2\leqslant\|\boldsymbol{x}\|^2+2\|\boldsymbol{x}\|\cdot\|\boldsymbol{y}\|+\|\boldsymbol{y}\|^2=(\|\boldsymbol{x}\|+\|\boldsymbol{y}\|)^2.$$

(iii)得到证明.

[1] 也称为柯西不等式(Cauchy inequality), 或柯西——许瓦尔兹不等式.

其中许瓦尔兹不等式(14)可以证明如下．因为，若 $\boldsymbol{y}=\mathbf{0}$，则(14)式成立，故假定 $\boldsymbol{y}\neq\mathbf{0}$．对于 $\forall\alpha\in F$，有

$$0\leqslant\langle\boldsymbol{x}-\alpha\boldsymbol{y},\ \boldsymbol{x}-\alpha\boldsymbol{y}\rangle=\langle\boldsymbol{x},\ \boldsymbol{x}\rangle-\alpha\langle\boldsymbol{y},\ \boldsymbol{x}\rangle-\bar{\alpha}\langle\boldsymbol{x},\ \boldsymbol{y}\rangle+|\alpha|^2\langle\boldsymbol{y},\ \boldsymbol{y}\rangle$$

若令式中 $\alpha=\langle\boldsymbol{x},\ \boldsymbol{y}\rangle/\langle\boldsymbol{y},\ \boldsymbol{y}\rangle$，则上式变成

$$0\leqslant\langle\boldsymbol{x},\ \boldsymbol{x}\rangle-|\langle\boldsymbol{x},\ \boldsymbol{y}\rangle^2/\langle\boldsymbol{y},\ \boldsymbol{y}\rangle.$$

这就意味着许瓦尔兹不等式成立.

试看一下起始讲过的向量空间 $C[t_0,\ t_1]^m$．对于 $\boldsymbol{f}\in C[t_0,\ t_1]^m$，由(4)式的内积导出的范数变成[1)]

$$\|\boldsymbol{f}\|=\langle\boldsymbol{f},\boldsymbol{f}\rangle^{1/2}\triangleq\left[\int_{t_0}^{t_1}\boldsymbol{f}^T(t)\boldsymbol{f}(t)\,dt\right]^{1/2}\tag{15}$$

［问题 6］ 当 $\boldsymbol{f}\in C[t_0,\ t_1]^m$ 时，$\int_{t_0}^{t_1}\boldsymbol{f}(t)\,dt\in R^m$，试证明

$$\left\|\int_{t_0}^{t_1}\boldsymbol{f}(t)\,dt\right\|_\infty\leqslant\int_{t_0}^{t_1}\|\boldsymbol{f}(t)\|_\infty\,dt$$

式中，$\|\cdot\|_\infty$ 是 R^m 中的范数，对于 $\boldsymbol{x}\in R^m$，设 $\|\boldsymbol{x}\|_\infty=\max\limits_i|x_i|$．对于范数，用 $\|\cdot\|_2$ 代替 $\|\cdot\|_\infty$，是否也有同样的结果.

1) t 固定，则 $\boldsymbol{f}(t)$ 是 R^m 中的向量，利用 R^m 中的范数 $\|\cdot\|_2$，得 $\boldsymbol{f}^T(t)\boldsymbol{f}(t)=\|\boldsymbol{f}(t)\|_2^2$．因此，(15)式可以写成 $\|\boldsymbol{f}\|=\left[\int_{t_0}^{t_1}\|\boldsymbol{f}(t)\|_2^2\,dt\right]^{1/2}$．这里要注意，$\|\boldsymbol{f}\|$ 是 $C[t_0,\ t_1]^m$ 的范数，$\|\boldsymbol{f}(t)\|_2$ 是 R^m 的范数．这样，在函数空间的情况，不会混淆.

当然，$C[t_0,\ t_1]^m$ 的范数不仅限于此．例如，当令 $\|\cdot\|$ 为 R^m 的任意范数时，$\|\boldsymbol{f}\|=\left[\int_{t_0}^{t_1}\|\boldsymbol{f}(t)\|^p\,dt\right]^{1/p}$，$1\leqslant p<\infty$ 变成 $C[t_0,\ t_1]^m$ 的范数.

B-I-10 伏龙斯基行列式及其推广

在 B-I-9 中讲过判断以分段连续函数为元素的一组向量是否线性独立的方法. 这里对所讨论的对象再稍加限制,假定只限于某些适当次连续可微的函数. 大家知道,关于这种函数的线性独立性,有下列性质.

设 $f_1(t), \cdots, f_n(t)$ 是在某区间 I 上定义的 $(n-1)$ 次连续可微的纯量实函数,则矩阵

$$\boldsymbol{W} \triangleq \begin{bmatrix} f_1 & f_1' & \cdots & f_1^{(n-1)} \\ f_2 & f_2' & \cdots & f_2^{(n-1)} \\ \vdots & \vdots & & \vdots \\ f_n & f_n' & \cdots & f_n^{(n-1)} \end{bmatrix}$$

的行列式,即 $\det \boldsymbol{W}$ 称为伏龙斯基行列式. 若在某个 $t \in I$, $\det \mathbf{W} \neq 0$, 则 $f_1(t), \cdots, f_n(t)$ 作为 t 的函数,是线性独立的.

这个结果可以推广到向量值函数的场合. 设 $\boldsymbol{f}_1(t), \cdots, \boldsymbol{f}_n(t)$ 是以在某个区间 I 上定义的 $(n-1)$ 次连续可微的实函数为元素的 r 维行向量,利用

$$\boldsymbol{F}(t) \triangleq \begin{bmatrix} \boldsymbol{f}_1(t) \\ \vdots \\ \boldsymbol{f}_n(t) \end{bmatrix} = n \times r$$

可对伏龙斯基矩阵 $\boldsymbol{W}$ 定义如下

$$\boldsymbol{W} \triangleq [\boldsymbol{F}(t) \vdots \boldsymbol{F}'(t) \vdots \cdots \vdots \boldsymbol{F}^{(n-1)}(t)] = n \times nr$$

[定理] 在区间 I 中的某个时刻 t', 若 $\boldsymbol{W}$ 的秩等于 n, 则向量 $\boldsymbol{f}_1(t), \cdots, \boldsymbol{f}_n(t)$ 在区间 I 上线性独立.

(证明) 若在区间 I 上向量线性相关, 则存在某个不为 $\mathbf{0}$ 的实向量 $\boldsymbol{\lambda} \in R^n$, 使得 $\boldsymbol{\lambda}^T \boldsymbol{F}(t) = \mathbf{0}$, $^\forall t \in I$ 成立. 可见, 若将该式对 t 微分 $(n-1)$ 次, 得 $\boldsymbol{\lambda}^T \boldsymbol{W} = \mathbf{0}\ ^\forall t \in I$. 这与在时刻 $t' \in I$, $\operatorname{rank} \boldsymbol{W} = n$ 相矛盾, 因此得证.

(注意)

(i) 该定理的逆一般不成立. 亦即, 对于所有的 $t \in I$, 仅仅 $\operatorname{rank} \boldsymbol{W} < n$, $\boldsymbol{f}_1(t), \cdots, \boldsymbol{f}_n(t)$ 还不一定线性相关.

(ii) 将函数的范围再加以限制, 只讨论区间 I 上的正则函数[1]. 那么, $\boldsymbol{f}_1(t), \cdots, \boldsymbol{f}_n(t)$ 线性独立的充分和必要条件是, 在任意 $t \in I$,

$$\operatorname{rank} [\boldsymbol{F}(t) \vdots \boldsymbol{F}'(t) \vdots \cdots \vdots \boldsymbol{F}^{(n-1)}(t) \vdots \cdots] = n$$

成立[23]. 证明省略.

---

1) 设 $f(\cdot)$ 为在实轴某开区间 I 上定义的实函数. 对于所有的 $t \in I$, 当 n 阶导函数 $f^{(n)}(\cdot)$ 连续存在时, 称为 $f(\cdot)$ 是 n 次连续可微或 C^n 级. 当 $f(\cdot)$ 是 C^∞ 级, 而且对于所有的 $t_0 \in I$, 存在某个正实数 ε, 在任意 $t \in (t_0 - \varepsilon, t_0 + \varepsilon)$, $f(t)$ 在 t_0 附近可以用泰勒级数展开

$$f(t) = \sum_{n=0}^{\infty} \frac{(t-t_0)^n}{n!} f^{(n)}(t_0)$$

表示时, 则称 $f(\cdot)$ 为正则函数 (regular function), 或称 $f(\cdot)$ 在 I 上是解析的 (analytic).

P-I-3 利普希茨条件

关于保证微分方程式解的存在及其唯一性的充分条件，有利普希茨条件(Lipschitz condition)．这一点在 A-I-1 中关于状态方程式部分已经谈到，现在再讨论一下这个条件．

所讨论的微分方程式是

$$\dot{\boldsymbol{x}}=\boldsymbol{f}(\boldsymbol{x},\ t),\ \boldsymbol{x}(t)\in R^n,\ \boldsymbol{f}:R^n\times R\to R^n \tag{1}$$

若将输入函数固定，作为已知函数处理，大家所熟悉的正规形状态方程式 $\dot{\boldsymbol{x}}=\tilde{\boldsymbol{f}}(\boldsymbol{x},\ \boldsymbol{u},\ t)$ 也可以变成(1)式形式．

(1)式形式的微分方程式，常常不一定具有唯一解，如已讲过，这一点由 $\dot{\boldsymbol{x}}=\sqrt{\boldsymbol{x}}$ 的例子可以理解．我们感兴趣的是，在(1)式中当给定任意起始时刻 t_0 和任意点 $\boldsymbol{x}_0\in R^n$ 时，对应于 $\boldsymbol{x}_0$，在 t_0 时刻通过 $\boldsymbol{x}_0$ 的解，在 $t\geqslant t_0$ 时存在且只有一个的条件．例如，方程式 $\dot{x}=x^2$，对应于起始条件 $t_0=0$，$x_0=1$，具有解 $x(t)=1/(1-t)$．但是，该解在 $t\to1$ 变成 $x(t)\to\infty$，不是在所有的 $t\geqslant0$ 都存在．也就是说，为了具有我们所希望的性质，可见(1)式右边的函数仅仅是对于 $\boldsymbol{x}$ 连续或可微还是不充分的．

从应用的角度来看，也要顺便考虑一下 $\boldsymbol{f}(\boldsymbol{x},\ t)$ 中将 $\boldsymbol{x}$ 固定，看作 t 的函数时某种程度的不连续性．否则，若所讨论的范围不包括阶跃输入等，对于自动控制中的应用是不方便的．

那么，考虑到以上各点的下列定理成立[28]．

［定理］ 在微分方程式 $\dot{\boldsymbol{x}}=\boldsymbol{f}(\boldsymbol{x},\ t)$ 中，将 $\boldsymbol{x}$ 固定在 R^n 的任意点时，设 $\boldsymbol{f}(\boldsymbol{x},\ \cdot)$ 是 t 的分段连续函数，而且可以找到这样的分段连续函数 $k(t)>0$，使得

$$\|\boldsymbol{f}(\boldsymbol{x},\ t)-\boldsymbol{f}(\boldsymbol{y},\ t)\|\leqslant k(t)\|\boldsymbol{x}-\boldsymbol{y}\| \tag{2}$$

$$\forall\boldsymbol{x},\ \boldsymbol{y}\in R^n,\ \forall t\in R$$

成立．式中 $\|\cdot\|$ 是 R^n 的(任意)范数．

这时，对于任意 $\boldsymbol{x}_0\in R^n$ 和任意 t_0，$\dot{\boldsymbol{x}}=\boldsymbol{f}(\boldsymbol{x},\ t)$ 的唯一解存在，该解满足起始条件 $\boldsymbol{x}(t_0)=\boldsymbol{x}_0$，对于 t 连续．

(注意)

(i) "有连续的唯一解存在"，详细地讲，就是下面的意思：有 t 的某个函数 $\boldsymbol{x}(t,\ \boldsymbol{x}_0)$ 存在，它满足 $\boldsymbol{x}(t_0,\ \boldsymbol{x}_0)=\boldsymbol{x}_0$，除了 $\boldsymbol{f}(\boldsymbol{x}(t,\ \boldsymbol{x}_0),\ t)$ 的不连续点外，对于任意 t，$\dot{\boldsymbol{x}}(t,\ \boldsymbol{x}_0)=\boldsymbol{f}(\boldsymbol{x}(t,\ \boldsymbol{x}_0),\ t)$ 均成立，而且这样的函数只有一个．

(ii) (2)式称为(大范围)利普希茨条件．以前讲过的方程式 $\dot{x}=x^2$，其右边函数是连续的，但是显而易见，没有满足条件"$|x^2-y^2|\leqslant k|x-y|$，$\forall x,\ y$"的常数 $k>0$ 存在．由此可见，$\boldsymbol{f}(\boldsymbol{x},\ t)$ 即使对于 $\boldsymbol{x}$ 连续，但不一定满足利普希茨条件．但是由(2)式可见，连续性是必要条件．在包含继电器的系统等当中，因为 $\boldsymbol{f}(\boldsymbol{x},\ t)$ 具有对于 $\boldsymbol{x}$ 不连续的特性，当然不符合利普希茨条件．在上述定理的范围内，不产生继电控制系统特有的跳跃现象．

(iii) 该定理表示，在 t_0 时刻通过 $\boldsymbol{x}_0$ 的解，不仅在 $t>t_0$，而且在过去的时间 $t<t_0$ 范

围内，也是唯一确定的．例如，$\dot{x}=x$, $x(0)=1$ 的解，不仅在 $t>0$，而且在 $t<0$ 也仅限于 $x(t)=e^t$.

不能按物理上时间进行的方向考虑(1)式微分方程式，而必须指定所指的仅仅是 $t>t_0$ 的解，或是还包括 $t<t_0$ 时的解．

自动控制中应用的线性时变系统　$\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}+\boldsymbol{B}(t)\boldsymbol{u}$; $\boldsymbol{y}=\boldsymbol{C}(t)\boldsymbol{x}+\boldsymbol{D}(t)\boldsymbol{u}$

现在，我们暂且来讲一下矩阵理论在上面标题所示系统分析(解的存在、解的各种性质等等)中的应用．

这种方程式出现在下列两种情况，① 所讨论的系统是由具有集中常数的线性变参数元素组合起来，模型化后得到的系统(例如，可参看本书第 90～92 页)；② 所讨论的系统可以用非线性状态方程式 $\dot{\boldsymbol{x}}=\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}, t)$ 表示，而在其某个特定解 $\tilde{\boldsymbol{x}}(t)=\boldsymbol{s}(t; \tilde{\boldsymbol{x}}(t_0), t_0, \tilde{\boldsymbol{u}})$ 的附近讨论系统的行为．因此，在应用上也是很重要的．

A-I-6 解的存在及其基本性质

解的存在和唯一性

首先，我们利用前一章的利普希茨条件来确定解的存在和唯一性.

[定理 1] 考虑线性变系数微分方程式

$$\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}+\boldsymbol{B}(t)\boldsymbol{u}(t). \tag{1}$$

设 $\boldsymbol{A}(\cdot)=n\times n$, $\boldsymbol{B}(\cdot)=n\times r$, $\boldsymbol{u}(\cdot)=r\times 1$, 都是以 t 的分段连续函数为元素的矩阵. 这时(1)式右边满足大范围利普希茨条件.

(问题 1) 试证明定理 1, 特别是要求出 P-I-3 中(2)式的 $k(t)$.

(提示): 利用 B-I-5 中矩阵范数.

根据定理 1, 若 $\boldsymbol{A}(\cdot)$, $\boldsymbol{B}(\cdot)$, $\boldsymbol{u}(\cdot)$的元素分段连续, 则可以保证在$(-\infty, \infty)$, (1)式的连续解的存在和唯一性. 而且, 在输出方程式

$$\boldsymbol{y}=\boldsymbol{C}(t)\boldsymbol{x}+\boldsymbol{D}(t)\boldsymbol{u}(t) \tag{2}$$

中, 设 $\boldsymbol{C}(\cdot)=m\times n$, $\boldsymbol{D}(\cdot)=m\times r$, 也都是以 t 的分段连续函数为元素的矩阵, 则对于$(-\infty, \infty)$, 输出 $\boldsymbol{y}$ 仍然是 t 的分段连续函数. 特别是, 即使不预先加以说明, 以下都认为上述假定成立.

当然, 这里解的唯一性, 是指当给定输入 $\boldsymbol{u}(\cdot)$时, 则满足起始条件的(1)式的解是唯一确定的. 若输入 $\boldsymbol{u}(\cdot)$不同, 解当然不同.

自由系统 $\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}$ 的解

因为即使令 $\boldsymbol{u}(\cdot)\equiv\boldsymbol{0}$, 定理 1 的前提条件当然还是成立, 所以仍然可以保证自由运动系统

$$\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x},\ \boldsymbol{x}(t_0)=\boldsymbol{x}_0 \tag{3}$$

的解存在和唯一性, 以及关于 t 的连续性. 我们将该解以 $\boldsymbol{x}_0$ 和 t_0 为参数, 写成 $\boldsymbol{\phi}(\cdot;\ \boldsymbol{x}_0,\ t_0)$, 再来讨论一下它的性质.

设 $\boldsymbol{x}(\cdot)\equiv\boldsymbol{0}$, 显然它是 $\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}$ 的解. 根据解的唯一性, 它是对于任意 $t_0\in R$, 满足 $\boldsymbol{x}(t_0)=\boldsymbol{0}$ 的唯一解. 亦即

[引理 1]

$$\boldsymbol{\phi}(\cdot;\ \boldsymbol{0},\ t_0)=\boldsymbol{0}\quad {}^{\forall}t_0\in R \tag{4}$$

成立.

[引理 2] $\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}$ 的解, 对于起始条件是线性的, 即

$$\boldsymbol{\phi}(\cdot;\ \alpha\boldsymbol{\xi}+\beta\boldsymbol{\eta},\ t_0)=\alpha\boldsymbol{\phi}(\cdot;\ \boldsymbol{\xi},\ t_0)+\beta\boldsymbol{\phi}(\cdot;\ \boldsymbol{\eta},\ t_0) \tag{5}$$

式中 $\alpha,\beta\in R$.

(证明) 将(5)式右边写成 $\bar{\phi}(\cdot)$, 则

$$\bar{\phi}(t_0)=\alpha\boldsymbol{\phi}(t_0;\ \boldsymbol{\xi},\ t_0)+\beta\boldsymbol{\phi}(t_0;\ \boldsymbol{\eta},\ t_0)=\alpha\boldsymbol{\xi}+\beta\boldsymbol{\eta},$$

而且

$$\dot{\bar{\phi}}(t)=\alpha\dot{\phi}(t;\ \boldsymbol{\xi},\ t_0)+\beta\dot{\phi}(t;\ \boldsymbol{\eta},\ t_0)$$
$$=\alpha\boldsymbol{A}(t)\boldsymbol{\phi}(t;\ \boldsymbol{\xi},\ t_0)+\beta\boldsymbol{A}(t)\boldsymbol{\phi}(t;\ \boldsymbol{\eta},\ t_0)$$
$$=\boldsymbol{A}(t)[\alpha\boldsymbol{\phi}(t;\ \boldsymbol{\xi},\ t_0)+\beta\boldsymbol{\phi}(t;\ \boldsymbol{\eta},\ t_0)=\boldsymbol{A}(t)\bar{\phi}(t)$$

根据解的唯一性，因而(5)式成立.

［定理 2］ $\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}$ 解的全体，形成实数体 R 上的向量空间.

［证明］ 只要证明 B-0-2 中定义 1 的 a)～f) 成立即可(试证明之).

该向量空间用 S 表示，S 称为解空间(Solution space)[1].

但是，在 S 的元 $\boldsymbol{\phi}(\cdot;\ \boldsymbol{x}_0,\ t_0)$中，若将变量 t 固定在某个值 $t_1\in R$，则 $\boldsymbol{\phi}(t_1;\ \boldsymbol{x}_0,\ t_0)$变成 R^n 的向量，而且有下列性质.

［引理 3］ 设 $\boldsymbol{a}_i\in R^n(i=1,\ \cdots,\ P)$，对于某个值 $t_1\in R$，若 $\boldsymbol{\phi}(t_1;\ \boldsymbol{a}_i,\ t_0)\ (i=1,\ \cdots,\ P)$线性独立(作为 R^n 的元)，则 $\boldsymbol{\phi}(\cdot;\ \boldsymbol{a}_i,\ t_0)\ (i=1,\ \cdots,\ P)$线性独立(作为 S 的元). 反之，若 $\boldsymbol{\phi}(\cdot;\ \boldsymbol{a}_i,\ t_0)$线性独立(作为 S 的元)，则对于任意 $t_1\in R$，$\boldsymbol{\phi}(t_1;\ \boldsymbol{a}_i,\ t_0)\ (i=1,\ \cdots,\ P)$线性独立[2](作为 R^n 的元)

(证明) 若 $\boldsymbol{\phi}(t_1;\ \boldsymbol{a}_i,\ t_0)$线性独立，则只有 $C_1=\cdots=C_P=0$ 时

$$\sum_{i=1}^{P}C_i\boldsymbol{\phi}(t_1;\ \boldsymbol{a}_i,\ t_0)=\boldsymbol{0}\tag{6}$$

才成立. 因此，也只有 $C_1=\cdots C_P=0$ 时

$$\sum_{i=1}^{P}C_i\boldsymbol{\phi}(t;\ \boldsymbol{a}_i,\ t_0)=\boldsymbol{0}\quad{}^{\forall}t\in R\tag{7}$$

才成立，则 $\boldsymbol{\phi}(\cdot;\ \boldsymbol{a}_i,\ t_0)$是线性独立的.

反之，设 $\boldsymbol{\phi}(\cdot;\ \boldsymbol{a}_i,\ t_0)$线性独立，而且对于某个 $t_1\in R$ 和一组不全为 0 的实数 $C_1,\ \cdots,\ C_P$，设

$$\sum_{i=1}^{P}C_i\boldsymbol{\phi}(t_1;\ \boldsymbol{a}_i,\ t_0)=\boldsymbol{0}\tag{8}$$

成立. 根据定理 2，$\bar{\boldsymbol{x}}(\cdot)\triangleq\sum_{i=1}^{P}C_i\boldsymbol{\phi}(\cdot;\ \boldsymbol{a}_i,\ t_0)$仍然是 $\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}$ 的一个解. 由(8)式及解的唯一性，得

$$\bar{\boldsymbol{x}}(\cdot)=\boldsymbol{\phi}(\cdot;\ \boldsymbol{0},\ t_1)\tag{9}$$

按着引理 1，

$$\bar{\boldsymbol{x}}(\cdot)=\sum_{i=1}^{P}C_i\boldsymbol{\phi}(\cdot;\ \boldsymbol{a}_i,\ t_0)=\boldsymbol{0}\tag{10}$$

这与 $\boldsymbol{\phi}(\cdot;\ \boldsymbol{a}_i,\ t_0)$的线性独立相矛盾. 证明完毕.

［系 1］ $\boldsymbol{a}_i(i=1,\ \cdots,\ P)$线性独立$\Leftrightarrow\boldsymbol{\phi}(\cdot;\ \boldsymbol{a}_i,\ t_0)\ (i=1,\ \cdots,\ P)$线性独立.

这里，作为起始条件 $\boldsymbol{x}(t_0)$，试给出向量 $\boldsymbol{e}_i$(参看 P-I-1).

［定理 3］ $\boldsymbol{\phi}(t;\ \boldsymbol{e}_i,\ t_0)\ (i=1,\ \cdots,\ n)$作为 S 的元是线性独立的，$\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}$ 的任意解

---

[1] 在$(-\infty,\ \infty)$上连续实函数的全体，是 R 上的向量空间(B-0-5). 若将其写成 $C(-\infty,\ \infty)$，根据定理 1，S 是$[C(-\infty,\ \infty)]^n$ 的部分空间.

[2] 注意，该性质仅限于解空间 S，所以对于$[C(-\infty,\ \infty)]^n$ 的全体不成立. 例如

$$\boldsymbol{x}_1(t)\triangleq\begin{bmatrix}t\\1\end{bmatrix}\quad\boldsymbol{x}_2(t)\triangleq\begin{bmatrix}1\\t^2\end{bmatrix}$$

作为$[C(-\infty,\ \infty)]^2$ 的元是线性独立的，但若固定在 $t=1$，所得到的 $\boldsymbol{x}_1(1)$ 和 $\boldsymbol{x}_2(1)$ 作为 R^2 的元，不是线性独立的.

可以用 $\boldsymbol{\phi}(t;\ \boldsymbol{e}_i,\ t_0)\ (i=1,\ \cdots,\ n)$ 的线性组合(唯一地)表示[1].

(证明) 由 $\boldsymbol{e}_i$ 的形式显而易见, $\boldsymbol{e}_i\ (i=1,\ \cdots,\ n)$ 作为 R^n 的元是线性独立的. 因此, 由引理 3, $\boldsymbol{\phi}(\cdot;\ \boldsymbol{e}_i,\ t_0)\ (i=1,\ \cdots,\ n)$ 是线性独立的.

考察 $\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}$ 的任意解 $\boldsymbol{x}(\cdot)$, 设其在 $t=t_0$ 时所取的值 $\boldsymbol{x}(t_0)$ 为 $\boldsymbol{a}\in R^n$, 显然

$$\boldsymbol{a}=\sum_{i=1}^{n} a_i\boldsymbol{e}_i \tag{11}$$

成立. 由引理 2 得

$$\boldsymbol{x}(\cdot)=\sum_{i=1}^{n} a_i\boldsymbol{\phi}(\cdot;\ \boldsymbol{e}_i,\ t_0). \tag{12}$$

证明完毕.

[系 2] 在 $\boldsymbol{\phi}(\cdot;\ \boldsymbol{e}_i,\ t_0)$ 中, 将 t 固定在任意 $t_1\in R$, 得到的 $\boldsymbol{\phi}(t_1;\ \boldsymbol{e}_i,\ t_0)\in R^n\ (i=1,\ \cdots,\ n)$, 作为 R^n 的向量是线性独立的.

状态转移矩阵

现在, 将 $\boldsymbol{\phi}(t;\ \boldsymbol{e}_i,\ t_0)$ 作为第 i 列, 组成矩阵 $\boldsymbol{\Phi}(t,\ t_0)$, 即

$$\boldsymbol{\Phi}(t,\ t_0)\triangleq[\boldsymbol{\phi}(t;\ \boldsymbol{e}_1,\ t_0)\cdots\boldsymbol{\phi}(t;\ \boldsymbol{e}_n,\ t_0)]=n\times n \tag{13}$$

由定理 3 的证明过程可见, 对于 $\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}$ 的任意解 $\boldsymbol{x}(\cdot)$,

$$\boldsymbol{x}(t)=\boldsymbol{\Phi}(t,\ t_0)\boldsymbol{x}(t_0)\quad {}^\forall t\in R \tag{14}$$

成立. 该式表明, 作用了矩阵 $\boldsymbol{\Phi}(t,\ t_0)$, 相当于使状态从 $\boldsymbol{x}(t_0)$ 转移到 $\boldsymbol{x}(t)$. 因此, 将 $\boldsymbol{\Phi}(\cdot,\ t_0)$ 称为状态转移矩阵(State transition matrix). 以上都把 t_0 看成某个固定的参数, 但是也可以将它看作在 $(-\infty,\ \infty)$ 范围内变化的变量, 这一点用文字 s 表示很清楚. 因此, $\boldsymbol{\Phi}(\cdot,\ \cdot)$ 变成两个变量的函数 $(R\times R\to R^{n\times n})$.

[定理 4] 状态转移矩阵 $\boldsymbol{\Phi}(t,\ s)$ 是矩阵微分方程式

$$\frac{\partial\boldsymbol{\Phi}(t,\ s)}{\partial t}=\boldsymbol{A}(t)\boldsymbol{\Phi}(t,\ s),\ \boldsymbol{\Phi}(s,\ s)=\boldsymbol{I}_n \tag{15}$$

的唯一解.

(问题 2) 试证明该定理.

(问题 3) 设

$$\boldsymbol{A}(t)=\begin{bmatrix}0 & \omega(t)\\ -\omega(t) & 0\end{bmatrix},$$

试证明状态转移矩阵变成

$$\boldsymbol{\Phi}(t,\ t_0)=\begin{bmatrix}\cos\Omega(t) & \sin\Omega(t)\\ -\sin\Omega(t) & \cos\Omega(t)\end{bmatrix}$$

式中 $\Omega(t)=\int_{t_0}^{t}\omega(t')\,dt'$.

矩阵微分方程式

$$\dot{\boldsymbol{X}}(t)=\boldsymbol{A}(t)\boldsymbol{X}(t)$$

---

[1] 在向量空间 V 中, 当某一组向量 $\boldsymbol{x}_i\ (i=1,\ \cdots,\ n)\in V$ 是线性独立的, 而且任意 $\boldsymbol{x}\in V$ 可以用 $\boldsymbol{x}_i$ 的线性组合表示时, $\boldsymbol{x}_i\ (i=1,\ \cdots,\ n)$ 称为 V 的基底(basis). 使用这个术语, [定理 3]可以表示成: $\boldsymbol{\phi}(\cdot;\ \boldsymbol{e}_i,\ t_0)\ (i=1,\ \cdots,\ n)$ 是向量空间 S 的一组基底.

的任意解(其中假定起始值 $\boldsymbol{X}(t_0)$ 是正则矩阵),称为 $\boldsymbol{A}(\cdot)$ 的基本矩阵(fundamental matrix).

(问题 4)

(i) 基本矩阵在任意 $t\in R$ 是正则的.

(ii) 设 $\boldsymbol{X}_1(\cdot)$, $\boldsymbol{X}_2(\cdot)$ 为 $\boldsymbol{A}(\cdot)$ 的任意两个基本矩阵, 则利用适当的正则矩阵 $\boldsymbol{T}$ 可以表示成

$$\boldsymbol{X}_1(t)=\boldsymbol{X}_2(t)\boldsymbol{T}. \quad \forall t\in R$$

(iii) 利用任意基本矩阵,可以将状态转移矩阵表示成

$$\boldsymbol{\Phi}(t,\ s)=\boldsymbol{X}(t)[\boldsymbol{X}(s)]^{-1}. \quad \forall t,\ s\in R \tag{16}$$

[定理 5] 关于 $\boldsymbol{A}(\cdot)$ 的状态转移矩阵有下列性质.

(i)
$$\boldsymbol{\Phi}(t,\ s)=\boldsymbol{\Phi}(t,\ t_1)\boldsymbol{\Phi}(t_1,\ s) \quad \forall t_1\in R \tag{17}$$

(ii) 逆矩阵存在

$$[\boldsymbol{\Phi}(t,\ s)]^{-1}=\boldsymbol{\Phi}(s,\ t) \tag{18}$$

(iii)
$$\frac{\partial}{\partial t}\boldsymbol{\Phi}(s,\ t)=-\boldsymbol{\Phi}(s,\ t)\boldsymbol{A}(t) \quad \boldsymbol{\Phi}(t,\ t)=\boldsymbol{I}_n \tag{19}$$

(iv)
$$\frac{\partial}{\partial t}\det\boldsymbol{\Phi}(t,\ s)=[\operatorname{tr}\boldsymbol{A}(t)]\det\boldsymbol{\Phi}(t,\ s) \tag{20}$$

(v) 设 $-\boldsymbol{A}^T(\cdot)$ 的状态转移矩阵为 $\overline{\psi}$, 则

$$\overline{\psi}(t,\ s)=[\boldsymbol{\Phi}^T(t,\ s)]^{-1} \tag{21}$$

(vi) 若有满足 $\|\boldsymbol{A}(t)\|\leqslant K \ \forall t\in R$ 的实数 K 存在,则

$$\|\boldsymbol{\Phi}(t,\ s)\|\leqslant\exp[K(t-s)]. \quad \forall t\geqslant s \tag{22}$$

式中 $\|\cdot\|$ 是由向量导出的矩阵范数.

(证明)

(i) 由(16)式,得

$$\boldsymbol{\Phi}(t,\ s)=\boldsymbol{X}(t)\boldsymbol{X}(s)^{-1}=[\boldsymbol{X}(t)\boldsymbol{X}(t_1)^{-1}][\boldsymbol{X}(t_1)\boldsymbol{X}(s)^{-1}]=\boldsymbol{\Phi}(t,\ t_1)\boldsymbol{\Phi}(t_1,\ s)^* \tag{23}$$

(ii) $\boldsymbol{\Phi}(t,\ s)\boldsymbol{\Phi}(s,\ t)=\boldsymbol{\Phi}(t,\ t)=\boldsymbol{I}$

$$\boldsymbol{\Phi}(s,\ t)\boldsymbol{\Phi}(t,\ s)=\boldsymbol{\Phi}(s,\ s)=\boldsymbol{I}$$

(iii) 由(18)式,得

$$\boldsymbol{\Phi}(s,\ t)=[\boldsymbol{\Phi}(t,\ s)]^{-1}$$

$$\frac{\partial}{\partial t}\boldsymbol{\Phi}(s,\ t)=\frac{\partial}{\partial t}[\boldsymbol{\Phi}(t,\ s)]^{-1}=-[\boldsymbol{\Phi}(t,\ s)]^{-1}\left[\frac{\partial}{\partial t}\boldsymbol{\Phi}(t,\ s)\right][\boldsymbol{\Phi}(t,\ s)]^{-1} \tag{24}$$

在这里,使用了 B-I-7 中的(9)式. 根据定理 4,

$$\frac{\partial}{\partial t}\boldsymbol{\Phi}(t,\ s)=\boldsymbol{A}(t)\boldsymbol{\Phi}(t,\ s) \tag{25}$$

将其代入(24)式,便得到(19)式.

(iv) 由 P-I-2 中(24)式,得

$$\frac{d}{dt}\det\boldsymbol{\Phi}(t,\ s)=\sum_{i=1}^{n}\det\boldsymbol{\Phi}^i(t,\ s) \tag{26}$$

* 原文误印为 $\cdots=\boldsymbol{\Phi}(t,\ t_1)\boldsymbol{\Phi}(t,\ s)$. ——译者注

式中 $\boldsymbol{\Phi}^i(t, s)$ 是用 $\boldsymbol{\Phi}(t, s)$ 的第 i 行对 t 的微分置换成的. 若将 $\boldsymbol{\Phi}(t, s)$ 的第 i 行写成 $\varphi_i(t, s)$, 则

$$\frac{\partial}{\partial t}\varphi_i(t, s)=\sum_{j=1}^{n}a_{ij}(t)\varphi_i(t_1, s) \tag{27}$$

利用 P-I-2 中行列式的性质(iii), (v), 得

$$\det\boldsymbol{\Phi}^i=a_{ii}(t)\det\boldsymbol{\Phi}(t, s) \tag{28}$$

由(26), (28)式可以证明(20)式.

(v) 根据定理 4,

$$\frac{\partial}{\partial t}\bar{\psi}(t, s)=-\boldsymbol{A}^T(t)\bar{\psi}(t, s),\ \bar{\psi}(s, s)=\boldsymbol{I}_n \tag{29}$$

此外, 将(19)式两边转置, 得

$$\frac{\partial}{\partial t}\boldsymbol{\Phi}^T(s, t)=-\boldsymbol{A}^T(t)\boldsymbol{\Phi}^T(s, t),\ \boldsymbol{\Phi}(s, s)=\boldsymbol{I}_n \tag{30}$$

若利用(19)式, 则

$$\frac{\partial}{\partial t}[\boldsymbol{\Phi}^T(t, s)]^{-1}=-\boldsymbol{A}^T(t)[\boldsymbol{\Phi}^T(t, s)]^{-1},\ \boldsymbol{\Phi}(s, s)=\boldsymbol{I}_n \tag{31}$$

比较(29), (31)式, 利用解的唯一性, 可以得到(21)式.

(vi) 开始先证明在给定的条件下

$$\|\boldsymbol{\phi}(t;\ \boldsymbol{x}_0,\ t_0)\|\leqslant\|\boldsymbol{x}_0\|\exp[K(t-t_0)],\ {}^\forall\boldsymbol{x}_0\in R^n,\ {}^\forall t\geqslant t_0 \tag{32}$$

首先, 因

$$\boldsymbol{\phi}(t;\ \boldsymbol{x}_0,\ t_0)=\boldsymbol{x}_0+\int_{t_0}^{t}\boldsymbol{A}(\tau)\boldsymbol{\phi}(\tau;\ \boldsymbol{x}_0,\ t_0)d\tau. \tag{33}$$

将上式两边取范数 $\|\cdot\|$, 利用 B-I-9 中(问题 6)的性质, 三角不等式及 $\|\boldsymbol{A}\boldsymbol{x}\|\leqslant\|\boldsymbol{A}\|\cdot\|\boldsymbol{x}\|$, 得

$$\|\boldsymbol{\phi}(t;\ \boldsymbol{x}_0,\ t_0)\|\leqslant\|\boldsymbol{x}_0\|+\int_{t_0}^{t}\|\boldsymbol{A}(\tau)\|\cdot\|\boldsymbol{\phi}(\tau;\ \boldsymbol{x}_0,\ t_0)\|d\tau. \tag{34}$$

因此, 由 Bellman-Gronwall 引理[1], 则(32)式成立.

那么, 根据定义,

$$\|\boldsymbol{\Phi}(t,\ t_0)\|=\max_{\|\boldsymbol{x}\|=1}\|\boldsymbol{\Phi}(t,\ t_0)\boldsymbol{x}\|=\max_{\|\boldsymbol{x}\|=1}\|\boldsymbol{\phi}(t;\ \boldsymbol{x},\ t_0)\|. \tag{35}$$

根据(32)式, 对于 $\|\boldsymbol{x}\|=1$ 的任意 $\boldsymbol{x}$, 则

$$\|\boldsymbol{\phi}(t;\ \boldsymbol{x},\ t_0)\|\leqslant\exp[K(t-t_0)] \tag{36}$$

由(35), (36)式可以证明(22)式. 证明完毕.

(注意)

由性质(iv), $\det\boldsymbol{\Phi}(t,\ t_0)$ 可以表示成

$$\det\boldsymbol{\Phi}(t,\ t_0)=\exp\int_{t_0}^{t}\operatorname{tr}\boldsymbol{A}(t')dt'. \tag{37}$$

---

[1] Bellman-Gronwall 的引理: 设 k, w 是在 $[t_0, \infty]$ 连续的函数, 而 $k\geqslant 0$, c 为常数. 若

$$w(t)\leqslant c+\int_{t_0}^{t}k(\tau)w(\tau)d\tau,\ {}^\forall t\geqslant t_0$$

则

$$w(t)\leqslant c\exp\int_{t_0}^{t}k(\tau)d\tau\quad {}^\forall t\geqslant t_0$$

成立.

系统 $\dot{\boldsymbol{\xi}}=-\boldsymbol{A}^T(t)\boldsymbol{\xi}$ 称为系统 $\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}$ 的伴随系统(adjoint system). 伴随系统的伴随系统，和原系统一致. 当伴随系统等于其原系统本身时，该系统称为是自伴的(Self adjoint). 系统自伴的充分和必要条件是，$\boldsymbol{A}(\cdot)$是斜对称矩阵.

(问题 5) 设 $\boldsymbol{x}(\cdot)$, $\boldsymbol{\xi}(\cdot)$分别为 $\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}$, $\dot{\boldsymbol{\xi}}=-\boldsymbol{A}^T\boldsymbol{\xi}$ 的解，试证明 $\boldsymbol{\xi}^T(t)\boldsymbol{x}(t)$ 是与 t 无关的常数.

有输入情况下的解

利用以上关于自由系统 $\dot{\boldsymbol{x}}(t)=\boldsymbol{A}(t)\boldsymbol{x}(t)$, $\boldsymbol{x}(t_0)=\boldsymbol{x}_0$ 讲过的结果，立即可求出加上输入后状态方程式和输出方程式的解. 首先，考虑

$$\dot{\boldsymbol{x}}(t)=\boldsymbol{A}(t)\boldsymbol{x}(t)+\boldsymbol{B}(t)\boldsymbol{u}(t),\ \boldsymbol{x}(t_0)=\boldsymbol{x}_0 \tag{1}$$

如以前讲过，这里假定 $\boldsymbol{A}(t)\in R^{n\times n}$, $\boldsymbol{B}(t)\in R^{n\times r}$, $\boldsymbol{x}(t)\in R^n$, $\boldsymbol{u}(t)\in R^r$, 其中 $\boldsymbol{A}(\cdot)$, $\boldsymbol{B}(\cdot)$, $\boldsymbol{u}(\cdot)$均对于 t 分段连续.

令 $\boldsymbol{z}(t)=\boldsymbol{\Phi}(t_0,\ t)\boldsymbol{x}(t)$，引入新的变量 $\boldsymbol{z}(t)$，式中 $\boldsymbol{\Phi}(t,\ t_0)$ 是自由系统 $\dot{\boldsymbol{x}}(t)=\boldsymbol{A}(t)\boldsymbol{x}$ 的转移矩阵. 利用(1)式计算 $\dot{\boldsymbol{z}}(t)$，得

$$\dot{\boldsymbol{z}}(t)=\left(\frac{\partial}{\partial t}\boldsymbol{\Phi}(t,\ t_0)+\boldsymbol{\Phi}(t_0,\ t)\boldsymbol{A}(t)\right)\boldsymbol{x}(t)+\boldsymbol{\Phi}(t_0,\ t)\boldsymbol{B}(t)\boldsymbol{u}(t)$$

用定理 5 的性质(iii)，消去上式右边第一项，得

$$\dot{\boldsymbol{z}}(t)=\boldsymbol{\Phi}(t_0,\ t)\boldsymbol{B}(t)\boldsymbol{u}(t). \tag{40}$$

(40)式可以直接积分，得

$$\boldsymbol{z}(t)=\boldsymbol{z}(t_0)+\int_{t_0}^{t}\boldsymbol{\Phi}(t_0,\ \tau)\boldsymbol{B}(\tau)\boldsymbol{u}(\tau)d\tau.$$

再代回原来的变量 $\boldsymbol{x}(t)=\boldsymbol{\Phi}(t_0,\ t)^{-1}\boldsymbol{z}(t)=\boldsymbol{\Phi}(t,\ t_0)\boldsymbol{z}(t)$，得

$$\boldsymbol{x}(t)=\boldsymbol{\Phi}(t,\ t_0)\left[\boldsymbol{x}_0+\int_{t_0}^{t}\boldsymbol{\Phi}(t_0,\ \tau)\boldsymbol{B}(\tau)\boldsymbol{u}(\tau)d\tau\right]. \tag{41}$$

为了弄清(14)式就是所要求的解，只要证明它在 $t=t_0$ 时的起始条件 $\boldsymbol{x}(t_0)=\boldsymbol{x}_0$ 成立及满足微分方程式(1)即可. 读者可以试一下.

将状态方程式(1)的解代入(2)式，便可得到输出方程式的解. 将以上结果归纳后如下.

[定理 6] 状态方程式 $\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}+\boldsymbol{B}(t)\boldsymbol{u}$, $\boldsymbol{x}(t_0)=\boldsymbol{x}_0$ 及输出方程式 $\boldsymbol{y}=\boldsymbol{C}(t)\boldsymbol{x}+\boldsymbol{D}(t)\boldsymbol{u}$ 的唯一解可由

$$\boldsymbol{x}(t)=\boldsymbol{\Phi}(t,\ t_0)\left[\boldsymbol{x}_0+\int_{t_0}^{t}\boldsymbol{\Phi}(t_0,\ \tau)\boldsymbol{B}(\tau)\boldsymbol{u}(\tau)d\tau\right] \tag{41}$$

$$\boldsymbol{y}(t)=\boldsymbol{C}(t)\boldsymbol{\Phi}(t,\ t_0)\left[\boldsymbol{x}_0+\int_{t_0}^{t}\boldsymbol{\Phi}(t_0,\ \tau)\boldsymbol{B}(\tau)\boldsymbol{u}(\tau)d\tau\right]+\boldsymbol{D}(t)\boldsymbol{u}(t) \tag{42}$$

给出.

利用转移矩阵的性质 $\boldsymbol{\Phi}(t,\ t_0)\boldsymbol{\Phi}(t_0,\ \tau)=\boldsymbol{\Phi}(t,\ \tau)$，及伴随系统的转移矩阵 $\boldsymbol{\Psi}(t,\ t_0)$ 的性质 $\boldsymbol{\Psi}(t,\ t_0)=\boldsymbol{\Phi}^T(t_0,\ t)$，则(41)式可等价地改写如下

$$\boldsymbol{x}(t)=\boldsymbol{\Phi}(t,\ t_0)\boldsymbol{x}_0+\int_{t_0}^{t}\boldsymbol{\Phi}(t,\ \tau)\boldsymbol{B}(\tau)\boldsymbol{u}(\tau)d\tau \tag{43}$$

$$=\boldsymbol{\Phi}(t,\ t_0)\boldsymbol{x}_0+\int_{t_0}^{t}\boldsymbol{\Psi}^T(\tau,\ t)\boldsymbol{B}(\tau)\boldsymbol{u}(\tau)d\tau \tag{44}$$

(41)式(或(43), (44)式)右边是规定状态的时间转移的式子. 显然, 它表示 $\boldsymbol{x}(t)$ 可以由 $\boldsymbol{x}_0$ 和区间 $[t_0, t]$ 上的输入 $\boldsymbol{u}[t_0, t]$ 决定. 亦即, 若利用 A-I-1 中(7)式所示的状态转移函数的符号, 则

$$\boldsymbol{x}(t)=\boldsymbol{S}(t;\ \boldsymbol{x}_0,\ t_0;\ \boldsymbol{u}[t_0,\ t])=\boldsymbol{\Phi}(t,\ t_0)\boldsymbol{x}_0+\int_{t_0}^{t}\boldsymbol{\Phi}(t,\ \tau)\boldsymbol{B}(\tau)\boldsymbol{u}(\tau)d\tau. \tag{45}$$

同样, (42)式右边是给出所讨论线性系统响应函数(A-I-1 中(9)式)的表示式. 即

$$\begin{aligned}\boldsymbol{y}(t)&=\boldsymbol{\rho}(t;\ \boldsymbol{x}_0,\ t_0;\ \boldsymbol{u}[t_0,\ t])\\&=\underbrace{\boldsymbol{C}(t)\boldsymbol{\Phi}(t,\ t_0)\boldsymbol{x}_0}_{\text{零输入响应}}+\underbrace{\boldsymbol{C}(t)\int_{t_0}^{t}\boldsymbol{\Phi}(t,\ \tau)\boldsymbol{B}(\tau)\boldsymbol{u}(\tau)d\tau+\boldsymbol{D}(t)\boldsymbol{u}(t)}_{\text{零状态响应}}.\end{aligned} \tag{46}$$

由(46)式可见, 输出 $\boldsymbol{y}(t)$ 可以表示成两个分量的和. 亦即, 右边第一项是, 假定 $\boldsymbol{u}(t)=\mathbf{0}$, $t\geqslant t_0$, 仅给出 x_0 时产生的输出分量, 叫做零输入响应(zero-input response); 第二项和第三项是, 假定 $\boldsymbol{x}_0=\mathbf{0}$, 仅给出输入时产生的输出分量. 因此叫做零状态响应(zero-state response). 由(46)式显而易见, 零输入响应及零状态响应分别为起始状态及输入的线性函数. 即

零输入响应的线性:

$$\boldsymbol{\rho}(t;\ \alpha\boldsymbol{x}_1+\beta\boldsymbol{x}_2,\ t_0;\ \mathbf{0})=\alpha\boldsymbol{\rho}(t;\ \boldsymbol{x}_1,\ t_0;\ \mathbf{0})+\beta\boldsymbol{\rho}(t;\ \boldsymbol{x}_2,\ t_0;\ \mathbf{0})$$
$$\forall\alpha、\beta\in R,\ \forall\boldsymbol{x}_1,\ \boldsymbol{x}_2\in R^n$$

零状态响应的线性:

$$\boldsymbol{\rho}(t;\ \mathbf{0},\ t_0;\ \alpha\boldsymbol{u}_1+\beta\boldsymbol{u}_2)=\alpha\boldsymbol{\rho}(t;\ \mathbf{0},\ t_0;\ \boldsymbol{u}_1)+\beta\boldsymbol{\rho}(t;\ \mathbf{0},\ t_0;\ \boldsymbol{u}_2)$$
$$\forall\alpha、\beta\in R,\ \boldsymbol{u}_1,\ \boldsymbol{u}_2=\text{分段连续的任意输入}$$

将该输入输出的性质归纳如下.

[系 1] 状态方程式用 $\dot{\boldsymbol{x}}(t)=\boldsymbol{A}(t)\boldsymbol{x}(t)+\boldsymbol{B}(t)\boldsymbol{u}(t)$, 输出方程式用 $\boldsymbol{y}(t)=\boldsymbol{C}(t)\boldsymbol{x}(t)+\boldsymbol{D}(t)\boldsymbol{u}(t)$ 表示的线性系统, 其输出是零输入响应及零状态响应之和; 零输入响应是起始状态的线性函数. 零状态响应是输入的线性函数.

(问题 6) 考虑矩阵微分方程式

$$\dot{\boldsymbol{X}}=\boldsymbol{A}(t)\boldsymbol{X}+\boldsymbol{X}\boldsymbol{B}(t)+\boldsymbol{C}(t)\qquad \boldsymbol{X}(t_0)=\boldsymbol{X}_0. \tag{47}$$

式中假定 $\boldsymbol{X}=n\times n$ 是未知矩阵, $\boldsymbol{A}(\cdot)$, $\boldsymbol{B}(\cdot)$, $\boldsymbol{C}(\cdot)$ 的元素是 t 的分段连续函数. 试证明该方程式解的存在和唯一性.

设 $\boldsymbol{\Phi}(\cdot,\ \cdot)$ 是对应于 $\boldsymbol{A}(\cdot)$ 的状态转移矩阵, $\boldsymbol{\Psi}(\cdot,\ \cdot)$ 是对应于 $\boldsymbol{B}^T(\cdot)$ 的状态转移矩阵, 试证明

$$X(t)=\boldsymbol{\Phi}(t,\ t_0)\boldsymbol{X}_0\boldsymbol{\Psi}^T(t,\ t_0)+\int_{t_0}^{t}\boldsymbol{\Phi}(t,\ \tau)\boldsymbol{C}(\tau)\boldsymbol{\Psi}^T(t,\ \tau)d\tau \tag{48}$$

成立.

提示: 前半部分, 将(47)式两边列展开, 得

$$\mathrm{cs}\,\dot{\boldsymbol{X}}=[(\boldsymbol{I}\otimes\boldsymbol{A}(t))+(\boldsymbol{B}^T(t)\otimes\boldsymbol{I})]\,\mathrm{cs}\,\boldsymbol{X}+\mathrm{cs}\,\boldsymbol{C}(t).$$

对它应用定理 1. 后半部分, 直接微分(48)式即可. 而且, 设

$$\boldsymbol{A}(t)\triangleq\boldsymbol{I}\otimes\boldsymbol{A}(t)+\boldsymbol{B}^T(t)\otimes\boldsymbol{I}.$$

证明 $\overline{\boldsymbol{A}}(\cdot)$ 的状态转移矩阵等于 $\boldsymbol{\Psi}(\cdot,\ \cdot)\otimes\boldsymbol{\Phi}(\cdot,\ \cdot)$ 即可.

A-I-7 输入输出关系

本章将假定起始状态 $\boldsymbol{x}_0$ 为 $\mathbf{0}$，专门研究系统输入 $\boldsymbol{u}$ 和输出 $\boldsymbol{y}$ 之间的关系.

脉冲响应

根据上一章(46)式，输出的零状态响应 $\boldsymbol{y}_{zs}$ 可以用

$$\boldsymbol{y}_{zs}(t)=\int_{t_0}^{t}\boldsymbol{C}(t)\boldsymbol{\Phi}(t,\ \tau)\boldsymbol{B}(\tau)\boldsymbol{u}(\tau)d\tau+\boldsymbol{D}(t)\boldsymbol{u}(t) \tag{1}$$

给出．若引入矩阵

$$\boldsymbol{H}(t,\ \tau)\triangleq\begin{cases}\boldsymbol{C}(t)\boldsymbol{\Phi}(t,\ \tau)\boldsymbol{B}(\tau)+\boldsymbol{D}(\tau)\delta(t-\tau) & t>\tau\\ \mathbf{0}, & t<\tau\end{cases} \tag{2}$$

式中 $\boldsymbol{H}(t,\ \tau)\in R^{m\times r}$，则(1)式可表示成[1)]

$$\boldsymbol{y}_{zs}(t)=\int_{t_0}^{t}\boldsymbol{H}(t,\ \tau)\boldsymbol{u}(\tau)d\tau\quad t>t_0. \tag{3}$$

其中 $\boldsymbol{H}(t,\ \tau)$ 的物理意义可以考虑如下.

设输入是在 $t=\tau$ 时刻加入的

$$\boldsymbol{u}(t)=\boldsymbol{e}_i\delta(t-\tau)$$

形式的脉冲，式中 $\boldsymbol{e}_i\in R^r$ 是第 i 个分量为 1，其它分量全为 0 的向量．对于该输入，零状态响应在 $t<\tau$ 时为 $\mathbf{0}$，当 $t>\tau$ 时为

$$\begin{aligned}\boldsymbol{y}_{zs}(t)&=\int_{t_0}^{t}\boldsymbol{C}(t)\boldsymbol{\Phi}(t,\ t')\boldsymbol{B}(t')\boldsymbol{e}_i\delta(t'-\tau)dt'+\boldsymbol{D}(t)\boldsymbol{e}_i\delta(t-\tau)\\&=[\boldsymbol{C}(t)\boldsymbol{\Phi}(t,\ \tau)\boldsymbol{B}(\tau)+\boldsymbol{D}(t)\delta(t-\tau)]\boldsymbol{e}_i\end{aligned}$$

即可以写成 $\boldsymbol{y}_{zs}(t)=\boldsymbol{H}(t,\ \tau)\boldsymbol{e}_i$．由此可见[2)]

$$\begin{aligned}\boldsymbol{H}(t,\ \tau)\boldsymbol{e}_i&=\boldsymbol{H}(t,\ \tau)\text{的第 } i \text{ 列}\\&=\text{所加输入为 }\boldsymbol{e}_i\delta(t-\tau)\text{时的零状态响应}(=\boldsymbol{\rho}(t;\ \mathbf{0},\ \tau_-;\ \boldsymbol{e}_i\delta(t-\tau)))\end{aligned} \tag{4}$$

因此，将(2)式定义的 $\boldsymbol{H}(t,\ \tau)$ 叫做用 $\dot{\boldsymbol{x}}(t)=\boldsymbol{A}(t)\boldsymbol{x}(t)+\boldsymbol{B}(t)\boldsymbol{u}(t)$，$\boldsymbol{y}(t)=\boldsymbol{C}(t)\boldsymbol{x}(t)+\boldsymbol{D}(t)\boldsymbol{u}(t)$表示的线性系统的脉冲响应(或脉冲响应矩阵 impulse response matrix).

（问题） 将前一章(1)，(2)的伴随系统定义为 $\dot{\boldsymbol{\xi}}=-\boldsymbol{A}^T(t)\boldsymbol{\xi}+\boldsymbol{C}^T(t)\boldsymbol{\gamma}$，$\boldsymbol{\eta}=\boldsymbol{B}^T(t)\boldsymbol{\xi}+\boldsymbol{D}^T(t)\boldsymbol{\gamma}$，设伴随系统的脉冲响应矩阵为 $\boldsymbol{K}(t,\ \tau)$．试证明 $\boldsymbol{K}(t,\ \tau)=[\boldsymbol{H}(\tau,\ t)]^T$.

代数等价系统

在求 $\dot{\boldsymbol{x}}(t)=\boldsymbol{A}(t)\boldsymbol{x}(t)+\boldsymbol{B}(t)\boldsymbol{u}(t)$，$\boldsymbol{x}(t_0)=\boldsymbol{x}_0$ 的解时已经看到，引入新变量 $\boldsymbol{z}(t)=\boldsymbol{\Phi}(t_0,\ t)\boldsymbol{x}(t)$后，关于 $\boldsymbol{z}$ 的微分方程式变得非常简单.

这里，从稍为一般性的情况出发，在按 $\tilde{\boldsymbol{x}}(t)=\boldsymbol{T}(t)\boldsymbol{x}(t)$引入新变量的情况下，考察关

---

1) $\boldsymbol{H}(t,\tau)$在 $t=\tau$ 产生的脉冲包含在积分范围，即积分上限可写成 t_+.

2) 在(4)式中，刚加脉冲之前，即在起始时刻 $t_0=\tau_-$ 时，假定起始状态为 $\mathbf{0}$．这是因为，在 $t=\tau$ 时刻脉冲一加上，一般会产生状态的不连续性($\boldsymbol{x}(\tau_+)\neq\boldsymbol{x}(\tau_-)$)，“$t_0=\tau$ 的起始状态”就没有意义.

于 $\tilde{\boldsymbol{x}}$ 成立的方程式和原方程式之间的关系．由此引入代数学上等价性的概念．现在假定，矩阵 $\boldsymbol{T}(t)\in R^{n\times n}$ 对于 t 连续；$\dot{\boldsymbol{T}}(t)$ 分段连续，$\det \boldsymbol{T}(t)\neq 0$，$^\forall t\in R$．由该变换矩阵 $\boldsymbol{T}(t)$ 可将 $\tilde{\boldsymbol{x}}$ 定义成

$$\tilde{\boldsymbol{x}}(t)=\boldsymbol{T}(t)\boldsymbol{x}(t). \tag{5}$$

容易证明，当对于 $\boldsymbol{x}$ 的状态方程式和输出方程式

$$\dot{\boldsymbol{x}}(t)=\boldsymbol{A}(t)\boldsymbol{x}(t)+\boldsymbol{B}(t)\boldsymbol{u}(t) \tag{6}$$

$$\boldsymbol{y}(t)=\boldsymbol{C}(t)\boldsymbol{x}(t)+\boldsymbol{D}(t)\boldsymbol{u}(t) \tag{7}$$

成立时，则对于新变量 $\tilde{\boldsymbol{x}}$，下列方程式成立

$$\dot{\tilde{\boldsymbol{x}}}(t)=\tilde{\boldsymbol{A}}(t)\tilde{\boldsymbol{x}}(t)+\tilde{\boldsymbol{B}}(t)\boldsymbol{u}(t) \tag{8}$$

$$\boldsymbol{y}(t)=\tilde{\boldsymbol{C}}(t)\tilde{\boldsymbol{x}}(t)+\tilde{\boldsymbol{D}}(t)\boldsymbol{u}(t) \tag{9}$$

式中

$$\left.\begin{aligned}\tilde{\boldsymbol{A}}(t)&=[\dot{\boldsymbol{T}}(t)+\boldsymbol{T}(t)\boldsymbol{A}(t)]\boldsymbol{T}^{-1}(t),\\ \tilde{\boldsymbol{B}}(t)&=\boldsymbol{T}(t)\boldsymbol{B}(t),\end{aligned}\right\} \tag{10}$$

$$\left.\begin{aligned}\tilde{\boldsymbol{C}}(t)&=\boldsymbol{C}(t)\boldsymbol{T}^{-1}(t),\\ \tilde{\boldsymbol{D}}(t)&=\boldsymbol{D}(t).\end{aligned}\right\} \tag{11}$$

一般，在用(6)，(7)式表示的线性系统和用(8)，(9)式表示的线性系统之间，当(10)，(11)式的关系式成立时(如上所述，假定 $\boldsymbol{T}(t)$ 是连续、正则的，$\dot{\boldsymbol{T}}(t)$ 分段连续)，称为两个系统是代数等价(algebraically equivalent)的[1]．

代数等价的两组线性系统，其转移矩阵、脉冲响应之间有着密切关系．首先，若利用(10)式，则可求出(8)，(9)式系统的转移矩阵 $\tilde{\boldsymbol{\Phi}}(t,\tau)$

$$\tilde{\boldsymbol{\Phi}}(t,\tau)=\boldsymbol{T}(t)\boldsymbol{\Phi}(t,\tau)\boldsymbol{T}^{-1}(\tau) \tag{12}$$

其原因是，它满足 $\tilde{\boldsymbol{\Phi}}(\tau,\tau)=\boldsymbol{I}$，$\dfrac{\partial}{\partial t}\tilde{\boldsymbol{\Phi}}(t,\tau)=\tilde{\boldsymbol{A}}(t)\tilde{\boldsymbol{\Phi}}(t,\tau)$．其次，若利用(12)式的结果，则可算出(8)，(9)式系统的脉冲响应

$$\begin{aligned}\tilde{\boldsymbol{H}}(t,\tau)&=\tilde{\boldsymbol{C}}(t)\tilde{\boldsymbol{\Phi}}(t,\tau)\tilde{\boldsymbol{B}}(\tau)+\tilde{\boldsymbol{D}}(t)\delta(t-\tau)\\ &=\boldsymbol{C}(t)\boldsymbol{T}^{-1}(t)\boldsymbol{T}(t)\boldsymbol{\Phi}(t,\tau)\boldsymbol{T}^{-1}(\tau)\boldsymbol{T}(\tau)\boldsymbol{B}(\tau)+\boldsymbol{D}(\tau)\delta(t-\tau)\\ &=\boldsymbol{H}(t,\tau)\quad t>\tau\end{aligned} \tag{13}$$

即可以看出，代数等价的两个系统有着相同的脉冲响应．这意味着，当加上同样输入时，两者的零状态响应完全一致．

实现问题

与以上讲过的内容相反，现在我们来讨论一下下列问题，即对于给定的脉冲响应矩阵

---

[1] 除了上述条件以外，对于某实数 k_1，k_2，当变换矩阵 $\boldsymbol{T}(t)$ 满足

$$\|\boldsymbol{T}(t)\|<k_1,\ \|\boldsymbol{T}^{-1}(t)\|<k_2$$

时，称为两个系统是拓扑等价(topologically equivalent)的，满足这种条件的 $\boldsymbol{T}(t)$ 称为李亚普诺夫变换(Lyapunov transformation)．考虑 $n=m=r=1$，$\boldsymbol{A}(t)=1$，$\boldsymbol{B}(t)=\boldsymbol{C}(t)=\boldsymbol{D}(t)=1$ 的系统，而且进行

$$\boldsymbol{T}(t)=e^t$$

的变换，则可得

$$\tilde{\boldsymbol{A}}(t)=2,\ \tilde{\boldsymbol{B}}(t)=e^t,\ \tilde{\boldsymbol{C}}(t)=e^{-t},\ \tilde{\boldsymbol{D}}(t)=1.$$

这两个系统是代数等价，而不是拓扑等价的．

$\boldsymbol{H}(t,\ \tau)$，求输入输出关系用 $\boldsymbol{H}(t,\ \tau)$ 表示的状态方程式 $\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}+\boldsymbol{B}(t)\boldsymbol{u}$，输出方程式 $\boldsymbol{y}=\boldsymbol{C}(t)\boldsymbol{x}+\boldsymbol{D}(t)\boldsymbol{u}$. 这叫做(脉冲响应的)实现问题(realization). 对于这个问题首先是，$\boldsymbol{H}(t,\ \tau)$具有什么形式才有实现的可能呢？即对于给定的 $\boldsymbol{H}(t,\ \tau)$，是否有满足(2)式的 $\boldsymbol{A}(t)$, $\boldsymbol{B}(t)$, $\boldsymbol{C}(t)$, $\boldsymbol{D}(t)$ 存在. 对此有下列定理.

[定理 1] $\boldsymbol{H}(t,\ \tau)$可实现的充分和必要条件是，利用以 t 的分段连续函数为元素的矩阵 $\boldsymbol{F}(\cdot)$, $\boldsymbol{G}(\cdot)$, $\boldsymbol{K}(\cdot)$，可以将它分解成

$$\boldsymbol{H}(t,\ \tau)=\boldsymbol{F}(t)\boldsymbol{G}(\tau)+\boldsymbol{K}(t)\delta(t-\tau)\quad {}^{\forall}t,\ \tau\in R,\ t>\tau. \tag{14}$$

(证明) 假定可以进行(14)式的分解，则

$$\boldsymbol{A}(t)=\boldsymbol{0},\ \boldsymbol{B}(t)=\boldsymbol{G}(t),\ \boldsymbol{C}(t)=\boldsymbol{F}(t),\ \boldsymbol{D}(t)=\boldsymbol{K}(t)\quad {}^{\forall}t\in \boldsymbol{R} \tag{15}$$

满足(2)式.

反之，若有满足(2)式的解存在，则利用基本矩阵可以分解成

$$\boldsymbol{H}(t,\ \tau)=[\boldsymbol{C}(t)\boldsymbol{X}(t)][\boldsymbol{X}^{-1}(\tau)\boldsymbol{B}(\tau)]+\boldsymbol{D}(t)\delta(t-\tau). \tag{16}$$

证明完毕

根据前节讲过的内容，若某个 $\boldsymbol{A}(t)$, $\boldsymbol{B}(t)$, $\boldsymbol{C}(t)$, $\boldsymbol{D}(t)$ 是 $\boldsymbol{H}(t,\ \tau)$ 的一个实现，则与其代数等价的系统都是 $\boldsymbol{H}(t,\ \tau)$ 的实现，即 $\boldsymbol{H}(t,\ \tau)$ 的实现不是唯一的. 其逆不成立，即具有相同脉冲响应矩阵的系统不一定是代数等价系统. 例如，

$$\begin{bmatrix}\dot{\boldsymbol{x}}_1\\ \dot{\boldsymbol{x}}_2\\ \dot{\boldsymbol{x}}_3\end{bmatrix}=\begin{bmatrix}\boldsymbol{A}_{11} & \boldsymbol{0} & \boldsymbol{A}_{13}\\ \boldsymbol{0} & \boldsymbol{A} & \boldsymbol{A}_{23}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{A}_{33}\end{bmatrix}\begin{bmatrix}\boldsymbol{x}_1\\ \boldsymbol{x}_2\\ \boldsymbol{x}_3\end{bmatrix}+\begin{bmatrix}\boldsymbol{B}_1\\ \boldsymbol{B}\\ \boldsymbol{0}\end{bmatrix}\boldsymbol{u} \tag{17}$$

$$\boldsymbol{y}=[\boldsymbol{0},\ \boldsymbol{C},\ \boldsymbol{C}_3]\begin{bmatrix}\boldsymbol{x}_1\\ \boldsymbol{x}_2\\ \boldsymbol{x}_3\end{bmatrix}+\boldsymbol{D}\boldsymbol{u} \tag{18}$$

具有和(6), (7)式系统相同的脉冲响应矩阵(由(17), (18)式的信号流图可见)，这里 n_1, n_2 是任意自然数，$\boldsymbol{A}_{11}(\cdot)=n_1\times n_1$, $\boldsymbol{A}_{13}(\cdot)=n_1\times n_3$, $\boldsymbol{A}_{23}(\cdot)=n_2\times n_3$, $\boldsymbol{A}_{33}(\cdot)=n_3\times n_3$, $\boldsymbol{B}_1(\cdot)=n_1\times r$, $\boldsymbol{C}_3(\cdot)=m\times n_3$ 是以 t 的分段连续函数为元素的任意矩阵.

在(17), (18)式中，$\bar{\boldsymbol{x}}\triangleq[\boldsymbol{x}_1^T\boldsymbol{x}_2^T\boldsymbol{x}_3^T]^T$ 是$(n+n_1+n_3)$维向量，而(6), (7)式的 $\boldsymbol{x}$ 是 n 维向量，不能象(5)式那样，用正则矩阵 $\boldsymbol{T}(t)$ 将 $\boldsymbol{x}$ 变换成 $\bar{\boldsymbol{x}}$. 因此，虽然这两个系统具有相同的脉冲响应，但不是代数等价的.

由上例可见，在所给 $\boldsymbol{H}(t,\ \tau)$的实现中，状态变量的个数 n 不是唯一的，可以多到任意个，但是最少个数有个限制. 一般，实现 $\boldsymbol{H}(t,\ \tau)$ 所必须的最少个数是确定的. 状态变量的个数 n 为最小时的实现称为最小实现(minimal realization).

A-I-8 可控性和可观测性

考虑线性变系数系统

$$\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}+\boldsymbol{B}(t)\boldsymbol{u} \tag{1}$$

$$\boldsymbol{y}=\boldsymbol{C}(t)\boldsymbol{x}+\boldsymbol{D}(t)\boldsymbol{u} \tag{2}$$

其中假定，$\boldsymbol{A}(\cdot)$，$\boldsymbol{B}(\cdot)$，$\boldsymbol{C}(\cdot)$，$\boldsymbol{D}(\cdot)$的元素是 t 的连续实函数．在 A-I-7 中，专门讨论了输入 $\boldsymbol{u}$ 和输出 $\boldsymbol{y}$ 之间的关系．本章将要讨论 $\boldsymbol{u}$、$\boldsymbol{y}$ 和系统内部状态 $\boldsymbol{x}$ 之间的关系．即讨论当加上输入 $\boldsymbol{u}$ 时，改变 $\boldsymbol{x}$ 可自由到什么程度，输出 $\boldsymbol{y}$ 仍然还能完全反映 $\boldsymbol{x}$ 的变化等问题．

如 A-I-6 中所述，在状态方程式(1)中，当输入 $\boldsymbol{u}$ 固定时，满足起始条件 $\boldsymbol{x}(t_0)=\boldsymbol{x}_0$ 的解仅有一个．但是对于不同的输入 $\boldsymbol{u}$，即使起始条件相同，也会得到不同的随时间变化的规律．因此将会产生相反的问题，假定满足方程式(1)的系统状态在 t_0 时刻为 $\boldsymbol{x}_0$，若恰当地选择 t_0 以后的输入 $\boldsymbol{u}$，是否能使未来的系统状态移到所希望的值 $\boldsymbol{x}_f$．当对这个问题的回答是肯定时，称为该系统是可控的．下面是更为严密的定义．

［定义 1］ 在状态方程式(1)中，设 $\boldsymbol{x}(t_0)=\boldsymbol{x}_0$．当在某个有限闭区间$[t_0,\ t_f]$加上适当的分段连续的输入 $\boldsymbol{u}_{[t_0,t_f]}$，可以使 $\boldsymbol{x}(t_f)=\boldsymbol{0}$ 时，称为状态 $\boldsymbol{x}_0$ 在 t_0 时刻可控(controllable)．当上述对任意 $\boldsymbol{x}_0\in R^n$ 均成立时，称为系统在 t_0 完全可控(Completely controllable)．而且，当对于任意 t_0 均成立时，称为系统是完全可控的[1]．

仍如 A-I-6 中所述，当状态 $\boldsymbol{x}(t)$ 和输入 $\boldsymbol{u}(t)$ 固定时，(2)式所给出的系统输出 $\boldsymbol{y}(t)$ 也是唯一确定的．那么反过来，若在某区间上知道输入 $\boldsymbol{u}$ 和输出 $\boldsymbol{y}$，对于这种输入，产生这种输出的状态 $\boldsymbol{x}$ 是否是唯一确定的．当回答是肯定时，称为该系统是可观测的．在介绍更严密的定义之前，我们来对可观测性作一些考察．首先，假设在(2)式中 $\boldsymbol{C}(\cdot)=n\times n$，$\det\boldsymbol{C}(t)\neq\boldsymbol{0}\ \forall t\in R$，显然可以解成

$$\boldsymbol{x}(t)=[\boldsymbol{C}(t)]^{-1}\boldsymbol{y}(t)-[\boldsymbol{C}(t)]^{-1}\boldsymbol{D}(t)\boldsymbol{u}(t). \tag{3}$$

$\boldsymbol{x}(t)$ 由 $\boldsymbol{y}(t)$，$\boldsymbol{u}(t)$ 唯一确定．除了这种情况，下面一般假定 $\boldsymbol{C}(\cdot)=m\times n(m\leqslant n)$．

如 A-I-6 中所述，(1)式的解为

$$\boldsymbol{x}(t)=\boldsymbol{\Phi}(t,\ t_0)\boldsymbol{x}_0+\int_{t_0}^{t}\boldsymbol{\Phi}(t,\ \tau)\boldsymbol{B}(\tau)\boldsymbol{u}(\tau)d\tau. \tag{4}$$

因为 $\boldsymbol{u}(\cdot)$ 作为已知，则上式右边第二项是已知量．因此，$\boldsymbol{x}(t)$ 唯一确定的充分必要条件是，右边第一项的 $\boldsymbol{x}_0$ 由 $\boldsymbol{u}(\cdot)$ 和 $\boldsymbol{y}(\cdot)$ 唯一确定．换句话说，有这样一个 t_f 存在．当输入 $\boldsymbol{u}(\cdot)$ 固定时，与两个以上不同起始状态相对应的输出 $\boldsymbol{y}[t_0,\ t_f]$ 完全不相同．因

$$\boldsymbol{y}(t)=\boldsymbol{C}(t)\boldsymbol{\Phi}(t,\ t_0)\boldsymbol{x}_0+\boldsymbol{C}(t)\int_{t_0}^{t}\boldsymbol{\Phi}(t,\ \tau)\boldsymbol{B}(\tau)\boldsymbol{u}(\tau)d\tau+\boldsymbol{D}(t)\boldsymbol{u}(t). \tag{5}$$

则上述成立的充分必要条件是，有这样一个 t_f 存在，当使 $\boldsymbol{u}(\cdot)\equiv\boldsymbol{0}$ 时，与不同起始状态相对应的输出在$[t_0,\ t_f]$上完全不相同．因此，以下假定 $\boldsymbol{u}(\cdot)\equiv\boldsymbol{0}$．首先，对于某个起始状态

[1] 在该定义中，t_f 没有指定．对此，若对任意 t_f 均为可控时，换句话说，当在包含在$[t_1,\ t_2]$内的任意闭区间均可控时，称为系统在 $[t_1,\ t_2]$ 上是总体可控的(totally controllable)[59]．

$\boldsymbol{x}_0 \neq \boldsymbol{0}$，若 $\boldsymbol{y}(\cdot) \equiv \boldsymbol{0}$，则不是可观测的．原因是，对于起始状态 $\boldsymbol{x}(t_0)=\boldsymbol{0}$，显然 $\boldsymbol{y}(\cdot) \equiv \boldsymbol{0}$．因为对于两个不同的起始状态可以得到同一个输出，则反而不是可观测的．也就是说，对于两个不同的起始状态 $\boldsymbol{x}_1$，$\boldsymbol{x}_2(\boldsymbol{x}_1 \neq \boldsymbol{x}_2)$，若假定可以得到同一个输出，则与 $\boldsymbol{x}_0 \triangleq \boldsymbol{x}_1-\boldsymbol{x}_2 \neq \boldsymbol{0}$ 的起始状态相对应的输出恒等于 $\boldsymbol{0}$．考虑到以上，作出如下定义．

［定义 2］ 在方程式

$$\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x},\ \boldsymbol{x}(t_0)=\boldsymbol{x}_0 \tag{6}$$

$$\boldsymbol{y}=\boldsymbol{C}(t)\boldsymbol{x} \tag{7}$$

中，对于 $\boldsymbol{x}_0 \neq \boldsymbol{0}$，当输出 $\boldsymbol{y}[t_0,\ t_f] \equiv \boldsymbol{0}$ 成立时，则称为状态 $\boldsymbol{x}_0$ 在区间 $[t_0,\ t_f]$ 上是不可观测的(unobservable)．若在时刻 t_0 选择适当的 t_f，使得不可观测的起始状态不存在时，称为系统在 t_0 是完全可观测的(completely observable)．当对于任意 t_0 都是完全可观测时，称为系统是完全可观测的．

由定义 1，2 可见，可控性与矩阵 $\boldsymbol{A}(\cdot)$，$\boldsymbol{B}(\cdot)$ 有关，可观测性与矩阵 $\boldsymbol{A}(\cdot)$，$\boldsymbol{C}(\cdot)$ 有关，而且二者均与矩阵 $\boldsymbol{D}(\cdot)$ 无关．因此，以下假定 $\boldsymbol{D}(\cdot) \equiv \boldsymbol{0}$．

（问题 1） 试证明，若系统在某区间 $[t_0,\ t_f]$ 上完全可控(完全可观测)，则在任意区间 $[t_0,\ t]$(其中 $t>t_f$)上是完全可控(完全可观测)的．

（问题 2） 设在 t_0 时刻可控的起始状态全体的集合为 C，试证明

若
$$\boldsymbol{x}_1 \in C,\ \boldsymbol{x}_2 \in C,$$
则
$$\alpha\boldsymbol{x}_1+\alpha_2\boldsymbol{x}_2 \in C,\ \forall\alpha_1,\ \alpha_2 \in R$$

成立．将上面的“可控”换成“不可观测”，试证明也可以得出同样结论．

于是，对于可控性可以得到下列充分必要条件．

［定理 1］ 系统在 t_0 时刻完全可控的充分必要条件是，对于某个 $t_f>t_0$，

$$\boldsymbol{W}_C(t_0,\ t_f) \triangleq \int_{t_0}^{t_f} \boldsymbol{\Phi}(t_0,\ t)\boldsymbol{B}(t)\boldsymbol{B}^T(t)\boldsymbol{\Phi}^T(t_0,\ t)dt \tag{8}$$

是正则矩阵[1)]．

由该定理很容易理解，系统完全可控的充分必要条件是，对于任意 t_0，上述条件是成立的．

$\boldsymbol{W}_C(t_0,\ t_f)$ 具有 B-I-9 中讲过的克兰姆矩阵的形状．因此也可以说，完全可控的充分必要条件是，对于某个 $t_f>t_0$，矩阵 $\boldsymbol{\Phi}(t_0,\ t)\boldsymbol{B}(t)$ 的行向量在区间 $[t_0,\ t_f]$ 上是线性独立的．

（定理 1 的证明） 对于某个 t_f，设 $\boldsymbol{W}_C(t_0,\ t_f)$ 是正则的．对于任意起始条件 $\boldsymbol{x}(t_0)=\boldsymbol{x}_0$，将输入 $\boldsymbol{u}(t)$ 选成

$$\boldsymbol{u}(t)=-\boldsymbol{B}^T(t)\boldsymbol{\Phi}^T(t_0,\ t)\boldsymbol{W}_C^{-1}(t_0,\ t_f)\boldsymbol{x}_0 \tag{9}$$

则由(4)式得

$$\begin{aligned}\boldsymbol{x}(t_f)&=\boldsymbol{\Phi}(t_f,\ t_0)\boldsymbol{x}_0-\int_{t_0}^{t_f}\boldsymbol{\Phi}(t_f,\ \tau)\boldsymbol{B}(\tau)\boldsymbol{B}^T(\tau)\boldsymbol{\Phi}^T(t_0,\ \tau)d\tau\,\boldsymbol{W}_C^{-1}(t_0,\ t_f)\boldsymbol{x}_0\\&=\boldsymbol{\Phi}(t_f,\ t_0)\boldsymbol{x}_0-\boldsymbol{\Phi}(t_f,\ t_0)\int_{t_0}^{t_f}\boldsymbol{\Phi}(t_0,\tau)\boldsymbol{B}(\tau)\boldsymbol{B}^T(\tau)\boldsymbol{\Phi}^T(t_0,\tau)d\tau\,\boldsymbol{W}_C^{-1}(t_0,\ t_f)\boldsymbol{x}_0\\&=\boldsymbol{\Phi}(t_f,\ t_0)\boldsymbol{x}_0-\boldsymbol{\Phi}(t_f,\ t_0)\boldsymbol{x}_0=\boldsymbol{0}.\end{aligned} \tag{10}$$

1) 当 $\boldsymbol{W}_C(t_0,\ t_f)$ 是正则矩阵时，由后面的证明也可以看到，这意味着，由 t_0 的任意起始状态 $\boldsymbol{x}(t_0)$ 出发，在 $t=t_f$ 时状态可以移到原点 $\boldsymbol{0}$．因此，当 $\det \boldsymbol{W}_C(t_0,\ t_f) \neq \boldsymbol{0}$ 时，也称为系统在区间 $[t_0,\ t_f]$ 上是可控的．

因为系统的状态可以从 $\boldsymbol{x}_0$ 移到 $\mathbf{0}$, 则系统是完全可控的.

反过来, 假定无论选择什么样的 $t_f>t_0$, $\boldsymbol{W}_C(t_0,\ t_f)$ 都不是正则矩阵, 则对于任意 $t_f>t_0$, 都存在有 $\boldsymbol{x}\neq\mathbf{0}$, 使得

$$\boldsymbol{x}^T\boldsymbol{W}_C(t_0,\ t_f)\boldsymbol{x}=\mathbf{0} \tag{11}$$

成立(参看 B-I-4). 若定义

$$\tilde{\boldsymbol{u}}(t)\triangleq-\boldsymbol{B}^T(t)\boldsymbol{\Phi}^T(t_0,\ t)\boldsymbol{x} \tag{12}$$

则

$$\int_{t_0}^{t_f}\tilde{\boldsymbol{u}}^T(t)\tilde{\boldsymbol{u}}(t)dt=\boldsymbol{x}^T\boldsymbol{W}_C(t_0,\ t_f)\boldsymbol{x}=\mathbf{0} \tag{13}$$

因 $\tilde{\boldsymbol{u}}(t)$ 是 t 的连续函数, 则由上式可见, 在区间 $[t_0,\ t_f]$ 上, $\tilde{\boldsymbol{u}}(\cdot)\equiv\mathbf{0}$.

但是, 若假定系统在 t_0 完全可控, 根据定义, 则对于某个时刻 t_f, 应该有满足

$$\boldsymbol{x}=-\int_{t_0}^{t_f}\boldsymbol{\Phi}(t_0,\ t)\boldsymbol{B}(t)\bar{\boldsymbol{u}}(t)dt \tag{14}$$

的 $\bar{\boldsymbol{u}}(t)$ 存在. 那么

$$\boldsymbol{x}^T\boldsymbol{x}=-\boldsymbol{x}^T\int_{t_0}^{t_f}\boldsymbol{\Phi}(t_0,\ t)\boldsymbol{B}(t)\bar{\boldsymbol{u}}(t)dt=-\int_{t_0}^{t}\tilde{\boldsymbol{u}}(t)\bar{\boldsymbol{u}}(t)dt=\mathbf{0}. \tag{15}$$

这与 $\boldsymbol{x}\neq\mathbf{0}$ 相矛盾. 证明完毕

将上面证明中(9)式的 $\bar{\boldsymbol{u}}(t)$ 稍微改变一下, 写成

$$\bar{\bar{\boldsymbol{u}}}(t)\triangleq-\boldsymbol{B}^T(t)\boldsymbol{\Phi}^T(t_0,\ t)\boldsymbol{W}_C^{-1}(t_0,\ t_f)\{\boldsymbol{x}_0-\boldsymbol{\Phi}(t_0,\ t_f)\boldsymbol{x}_f\} \tag{16}$$

式中 $\boldsymbol{x}_0$ 是给定的起始状态, $\boldsymbol{x}_f$ 是 R^n 的任意向量. 容易证明, 若加上该输入, 则系统状态在 $t=t_f$ 移到 $\boldsymbol{x}_f$. 这意味着, 若系统在定义 1 的意义上完全可控(即有从任意 $\boldsymbol{x}_0\in R^n$ 移到原点 $\mathbf{0}$ 的控制 $\boldsymbol{u}(\cdot)$ 存在), 实际上则有从任意 $\boldsymbol{x}_0\in R^n$ 移到 $\boldsymbol{x}_f\in R^n$ 的控制存在.

(问题 3) 试证明矩阵 $\boldsymbol{W}_C(t_0,\ t_f)$ 有下列性质.

(i) $\boldsymbol{W}_C(t_0,\ t_f)$ 是对称矩阵.

(ii) $\boldsymbol{W}_C(t_0,\ t_f)$ 是准正定的.

(iii) $$\frac{d}{dt}\boldsymbol{W}_C(t,\ t_f)=\boldsymbol{A}(t)\boldsymbol{W}_C(t,\ t_f)+\boldsymbol{W}_C(t,\ t_f)\boldsymbol{A}^T(t)-\boldsymbol{B}(t)\boldsymbol{B}^T(t)$$
$$\boldsymbol{W}_C(t_f,\ t_f)=\mathbf{0}$$

(iv) $\boldsymbol{W}_C(t_0,\ t_f)=\boldsymbol{W}_C(t_0,\ t)+\boldsymbol{\Phi}(t_0,\ t)\boldsymbol{W}_C(t,\ t_f)\boldsymbol{\Phi}^T(t_0,\ t)$

对于可观测性, 有下列与定理 1 很相似的性质.

[定理 2] 系统在 t_0 时刻完全可观测的充分必要条件是, 对于某个 $t_f>t_0$,

$$\boldsymbol{W}_0(t_0,\ t_f)\triangleq\int_{t_0}^{t_f}\boldsymbol{\Phi}^T(t,\ t_0)\boldsymbol{C}^T(t)\boldsymbol{C}(t)\boldsymbol{\Phi}(t,\ t_0)dt \tag{17}$$

是正则矩阵.

(证明) 设 $\boldsymbol{W}_0(t_0,\ t_f)$ 是正则矩阵. 一般, 零输入下的输出 $\boldsymbol{y}(t)$ 可以表示成

$$\boldsymbol{y}(t)=\boldsymbol{C}(t)\boldsymbol{\Phi}(t,\ t_0)\boldsymbol{x}_0 \tag{18}$$

将该式左乘以 $\boldsymbol{\Phi}^T(t,\ t_0)\boldsymbol{C}^T(t)$, 在 $[t_0,\ t_f]$ 上积分, 得

$$\int_{t_0}^{t_f}\boldsymbol{\Phi}^T(t,\ t_0)\boldsymbol{C}^T(t)\boldsymbol{y}(t)dt=\boldsymbol{W}_0(t_0,\ t_f)\boldsymbol{x}_0. \tag{19}$$

因为 $\boldsymbol{W}_0(t_0,\ t_f)$ 是正则矩阵, 则满足 $\boldsymbol{W}_0(t_0,\ t_f)\boldsymbol{x}_0=\mathbf{0}$ 的 $\boldsymbol{x}_0$ 只有 $\mathbf{0}$. 因此, 是完全可观测的. 实际上, 因为 $\boldsymbol{W}_0(t_0,\ t_f)$ 是正则矩阵, 则 $[\boldsymbol{W}_0(t_0,\ t_f)]^{-1}$ 存在, $\boldsymbol{x}_0$ 可由 $\boldsymbol{y}[t_0,\ t_f]$ 唯一地求出

$$\boldsymbol{x}_0=[\boldsymbol{W}_0(t_0,\ t_f)]^{-1}\int_{t_0}^{t_f}\boldsymbol{\Phi}^T(t_0,\ t)\boldsymbol{C}^T(t)\boldsymbol{y}(t)dt. \tag{20}$$

反之，假定无论选择什么样的 $t_f>t_0$，$\boldsymbol{W}_0(t_0,\ t_f)$ 都不是正则矩阵，则对于任意 $t_f>t_0$，有使

$$\boldsymbol{x}^T\boldsymbol{W}_0(t_0,\ t_f)\boldsymbol{x}=\boldsymbol{0} \tag{21}$$

成立的 $\boldsymbol{x}\neq\boldsymbol{0}$ 存在．该式可详细写成

$$\begin{aligned}\boldsymbol{x}^T\boldsymbol{W}_0(t_0,\ t_f)\boldsymbol{x}&=\boldsymbol{x}^T\int_{t_0}^{t_f}\boldsymbol{\Phi}^T(t,\ t_0)\boldsymbol{C}^T(t)\boldsymbol{C}(t)\boldsymbol{\Phi}(t,\ t_0)dt\,\boldsymbol{x}\\&=\int_{t_0}^{t_f}\{\boldsymbol{x}^T\boldsymbol{\Phi}^T(t,\ t_0)\boldsymbol{C}^T(t)\}\{\boldsymbol{C}(t)\boldsymbol{\Phi}(t,\ t_0)\boldsymbol{x}\}dt=0.\end{aligned} \tag{22}$$

因为 $\boldsymbol{C}(t)\boldsymbol{\Phi}(t,\ t_0)\boldsymbol{x}$ 是连续函数，则在 $[t_0,\ t_f]$ 的全区间

$$\boldsymbol{C}(t)\boldsymbol{\Phi}(t,\ t_0)\boldsymbol{x}=\boldsymbol{0} \tag{23}$$

成立．这意味着，以 $\boldsymbol{x}\neq\boldsymbol{0}$ 为起始状态的(6)，(7)式系统的输出在 $[t_0,\ t_f]$ 是 $\boldsymbol{0}$，系统不是可观测的．

证明完毕

现在对于系统

$$\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}+\boldsymbol{B}(t)\boldsymbol{u}, \tag{24}$$

$$\boldsymbol{y}=\boldsymbol{C}(t)\boldsymbol{x} \tag{25}$$

试组成另外一个系统

$$\dot{\boldsymbol{x}}^*=-\boldsymbol{A}^T(t)\boldsymbol{x}^*-\boldsymbol{C}^T(t)\boldsymbol{u}^* \tag{26}$$

$$\boldsymbol{y}^*=\boldsymbol{B}^T(t)\boldsymbol{x}^*. \tag{27}$$

由定理 1, 2 立即得下列系 1.

［系 1］ 系统(24)，(25)完全可控(完全可观测)的充分必要条件是，系统(26)，(27)是完全可观测(完全可控)的．

这个关系称为可控性和可观测性之间的对偶性(duality)，系统(26)，(27)称为(24)，(25)的对偶系统．

(问题 4) 试证明系 1.

特殊的变系数线性系统

以上所讨论的系统，其系数一般都是时间 t 的连续函数．若讨论的范围这么宽，下面关于常系数线性系统将要讲的许多性质，原封不动地拿到变系数系统是不成立的．但是，即使是变系数系统，如果仅仅以加上某些限制条件的一类系统为对象，则常系数系统所具有的某些性质，对于这样的系统通常也照样成立．因为这些限制条件多半与可控性、可观测性有关，所以这里要稍微讲一下．

开始先假定 $\boldsymbol{A}(t)$，$\boldsymbol{B}(t)$ 只是必要次连续可微，试将矩阵 $\boldsymbol{\Phi}(t_0,\ t)\boldsymbol{B}(t)$ 对 t 微分，由 A-I-6 中定理 5 之(iii)，得

$$\begin{aligned}\frac{\partial}{\partial t}\boldsymbol{\Phi}(t_0,\ t)\boldsymbol{B}(t)&=-\boldsymbol{\Phi}(t_0,\ t)\boldsymbol{A}(t)\boldsymbol{B}(t)+\boldsymbol{\Phi}(t_0,\ t)\frac{d\boldsymbol{B}(t)}{dt}\\&=\boldsymbol{\Phi}(t_0,\ t)\left\{-\boldsymbol{A}(t)\boldsymbol{B}(t)+\frac{d\boldsymbol{B}(t)}{dt}\right\}\end{aligned} \tag{28}$$

设最右边括号中为 $\boldsymbol{P}_1(t)$，则

$$\frac{\partial^2}{\partial t^2}\boldsymbol{\Phi}(t_0,\ t)\boldsymbol{B}(t)=\frac{\partial}{\partial t}\{\boldsymbol{\Phi}(t_0,\ t)\boldsymbol{P}_1(t)\}$$

$$=\boldsymbol{\Phi}(t_0,\ t)\left\{-\boldsymbol{A}(t)\boldsymbol{P}_1(t)+\frac{d\boldsymbol{P}_1(t)}{dt}\right\} \tag{29}$$

以下同样,可见一般微分 α 次最后得

$$\frac{\partial^\alpha}{\partial t^\alpha}\boldsymbol{\Phi}(t_0,\ t)\boldsymbol{B}(t)=\boldsymbol{\Phi}(t_0,\ t)\boldsymbol{P}_\alpha(t) \tag{30}$$

式中 $\boldsymbol{P}_j(t)=-\boldsymbol{A}(t)\boldsymbol{P}_{j-1}(t)+\dot{\boldsymbol{P}}_{j-1}(t)$, $\boldsymbol{P}_0(t)=\boldsymbol{B}(t)$. 而且立即可以看到,$\boldsymbol{\Phi}(t_0,\ t)\boldsymbol{B}(t)$ α 次连续可微的充分必要条件是, $\boldsymbol{A}(t)$ 为 $(\alpha-1)$ 次, $\boldsymbol{B}(t)$ 为 α 次连续可微.

将 $\boldsymbol{P}_j(t)$ 排列成矩阵

$$\boldsymbol{Q}_k(t)\triangleq[\boldsymbol{P}_0(t),\ \boldsymbol{P}_1(t),\ \cdots,\ \boldsymbol{P}_{k-1}(t)] \tag{31}$$

容易证明,该矩阵和可控性有关.

［系 2］ 系统在区间 $[t_0,\ t_f]$ 上完全可控的充分条件是,对于某个 $t\in[t_0,\ t_f]$,

$$\operatorname{rank}\boldsymbol{Q}_n(t)=n \tag{32}$$

成立[1].

（证明） 如定理 1 之后所述, 可控性的充分必要条件是, $\boldsymbol{\Phi}(t_0,\ t)\boldsymbol{B}(t)$ 的行向量在 $[t_0,\ t_f]$ 上是线性独立的. 根据 B-I-10 中的定理,其充分条件是,伏龙斯基矩阵

$$\boldsymbol{W}\triangleq\left[\boldsymbol{\Phi}(t_0,\ t)\boldsymbol{B}(t),\ \frac{\partial}{\partial t}\boldsymbol{\Phi}(t_0,\ t)\boldsymbol{B}(t),\ \cdots,\ \frac{\partial^{n-1}}{\partial t^{n-1}}\boldsymbol{\Phi}(t_0,\ t)\boldsymbol{B}(t)\right] \tag{33}$$

的秩在某个 $t\in[t_0,\ t_f]$ 上为 n. 显然

$$\boldsymbol{W}=\boldsymbol{\Phi}(t_0,\ t)\boldsymbol{Q}_n(t). \tag{34}$$

因 $\boldsymbol{\Phi}(t_0,\ t)$ 是正则矩阵,则 $\boldsymbol{W}$ 的秩等于 $\boldsymbol{Q}_n(t)$ 的秩.

证明完毕

同样,用迭代公式

$$\boldsymbol{S}_j(t)=\boldsymbol{S}_{j-1}(t)\boldsymbol{A}(t)+\dot{\boldsymbol{S}}_{j-1}(t),\ \boldsymbol{S}_0(t)=\boldsymbol{C}(t) \tag{35}$$

作矩阵 $\boldsymbol{S}_j(t)$, 定义

$$\boldsymbol{R}_k(t)\triangleq\begin{bmatrix}\boldsymbol{S}_0(t)\\ \boldsymbol{S}_1(t)\\ \vdots\\ \boldsymbol{S}_{k-1}(t)\end{bmatrix}. \tag{36}$$

（问题 5） 试叙述 $\boldsymbol{R}_k(t)$ 和可观测性的关系.

利用 $\boldsymbol{Q}_k(t)$, $\boldsymbol{R}_k(t)$, 定义定秩系统如下.

［定义 3］ 在系统

$$\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}+\boldsymbol{B}(t)\boldsymbol{u} \tag{37}$$

$$\boldsymbol{y}=\boldsymbol{C}(t)\boldsymbol{x} \tag{38}$$

中, 当有满足下列条件的正整数 α, β, q_c, q_0 存在时, 称为定秩系统 (constant rank system).

（i） $\boldsymbol{A}(t)$, $\boldsymbol{B}(t)$, $\boldsymbol{C}(t)$ 分别为 $\max(\alpha,\ \beta)-1$, α, β 次连续可微.

---

[1] 一般, 该条件不是必要的. 但是若限定 $\boldsymbol{A}(t)$, $\boldsymbol{B}(t)$ 为正则函数,则上述条件就变成在区间 $[t_0,\ t_f]$ 上可控的充分必要条件. 而且它同时也是在区间 $[t_0,\ t_f]$ 上总体可控 (totally controllable) 的充分必要条件.

(ii) $\operatorname{rank} \boldsymbol{Q}_\alpha(t) = \operatorname{rank} \boldsymbol{Q}_{\alpha+1}(t) = q_c \leqslant n \quad \forall t \in R$

$\operatorname{rank} \boldsymbol{R}_\beta(t) = \operatorname{rank} \boldsymbol{R}_{\beta+1}(t) = q_0 \leqslant n \quad \forall t \in R$

定秩系统具有什么性质，以后将要叙述．这里只是附带说一下，若在$(-\infty, \infty)$ $\boldsymbol{A}(t)$, $\boldsymbol{B}(t)$, $\boldsymbol{C}(t)$是正则函数，则为定秩系统(可以利用 B-I-10 中注意(ii)加以证明)．

下面，作为比可控性更强的条件，我们来定义一致可控性．以下关于对称矩阵$\boldsymbol{M}$, $\boldsymbol{N}$，将$(\boldsymbol{M}-\boldsymbol{N})$是正定(准正定)写成$\boldsymbol{M}>\boldsymbol{N}$($\boldsymbol{M}\geqslant\boldsymbol{N}$)．

[定义 4]　对于某实数σ及与σ有关的α_0, α_1, β_0, β_1, 当

(i) $0<\alpha_0(\sigma)\boldsymbol{I}\leqslant \boldsymbol{W}_C(t, t+\sigma)\leqslant\alpha_1(\sigma)\boldsymbol{I} \quad \forall t\in R$

(ii) $0<\beta_0(\sigma)\boldsymbol{I}\leqslant\boldsymbol{\Phi}(t+\sigma, t)\boldsymbol{W}_C(t, t+\sigma)\boldsymbol{\Phi}^T(t+\sigma, t)\leqslant\beta_1(\sigma)\boldsymbol{I} \quad \forall t\in R$

成立时，称为系统是一致完全可控的(uniformly completely controllable)．

同样

[定义 5]　对于某实数σ及与σ有关的α_0', α_1', β_0', β_1', 当

(i) $0<\alpha_0'(\sigma)\boldsymbol{I}\leqslant\boldsymbol{W}_0(t, t+\sigma)\leqslant\alpha_1'(\sigma)\boldsymbol{I} \quad \forall t\in R$

(ii) $0<\beta_0'(\sigma)\boldsymbol{I}\leqslant\boldsymbol{\Phi}^T(t, t+\sigma)\boldsymbol{W}_0(t, t+\sigma)\boldsymbol{\Phi}(t, t+\sigma)\leqslant\beta_1'(\sigma)\boldsymbol{I} \quad \forall t\in R$

成立时，称为系统是一致完全可观测的(uniformly completely observable)．

以上是一般的定义，而在系数矩阵范数有界的情况下，变得稍微简单一些．

[引理 1]　若有满足$\|\boldsymbol{A}(t)\|<K$, $\|\boldsymbol{B}(t)\|<K$, $\forall t\in R$的常数K存在，则系统一致完全可控的充分必要条件是，对于某个σ有使

$$0<\alpha_0(\alpha)\boldsymbol{I}<\boldsymbol{W}_C(t, t+\sigma) \quad \forall t\in R \tag{39}$$

成立的$\alpha_0(\sigma)$存在．

(证明)　必要性是显然的，下面只证明其充分性．亦即，若满足上面条件，只要证明它也满足定义 4 中余下的条件即可．首先

$$\|\boldsymbol{W}_C(t, t+\sigma)\|\leqslant\int_t^{t+\sigma}\|\boldsymbol{\Phi}(t, \tau)\|^2\|\boldsymbol{B}(\tau)\|^2 dt \tag{40}$$

因$\|\boldsymbol{B}(t)\|<K$，并且如 A-I-6 中所述，若$\|\boldsymbol{A}(t)\|<K$，则$\|\boldsymbol{\Phi}(t, \tau)\|<\exp K(t-\tau)$成立，于是有使

$$\|\boldsymbol{W}_C(t, t+\sigma)\|\leqslant\alpha_1(\sigma) \tag{41}$$

成立的$\alpha_1(\sigma)$存在．因此

$$\boldsymbol{W}_C(t, t+\sigma)\leqslant\alpha_1(\sigma)\boldsymbol{I} \tag{42}$$

因为$\boldsymbol{\Phi}(t+\sigma, t)$是正则矩阵，则

$$0<\beta_0(\sigma)\boldsymbol{I}\leqslant\boldsymbol{\Phi}(t+\sigma, t)\boldsymbol{W}_C(t, t+\sigma)\boldsymbol{\Phi}^T(t+\sigma, t). \tag{43}$$

而且

$$\|\boldsymbol{\Phi}(t+\sigma, t)\boldsymbol{W}_C(t, t+\sigma)\boldsymbol{\Phi}^T(t+\sigma, t)\|\leqslant\|\boldsymbol{\Phi}(t+\sigma, t)\|^2\|\boldsymbol{W}_C(t, t+\sigma)\| \tag{44}$$

和前面一样，考虑到(41)式，可见有使

$$\|\boldsymbol{\Phi}(t+\sigma, t)\boldsymbol{W}_C(t, t+\sigma)\boldsymbol{\Phi}^T(t+\sigma, t)\|\leqslant\beta_1(\sigma) \tag{45}$$

成立的$\beta_1(\sigma)$存在．因此

$$\boldsymbol{\Phi}(t+\sigma, t)\boldsymbol{W}_C(t, t+\sigma)\boldsymbol{\Phi}^T(t+\sigma, t)\leqslant\beta_1(\alpha)\boldsymbol{I}. \tag{46}$$

证明完毕

当然,对于一致完全可观测性也可以得到同样的结果. 基于这些结论,一致系统可定义如下.

[定义 6] 满足下列条件的系统称为一致系统(uniform system).

(i) $\|\boldsymbol{A}(t)\|<K$, $\|\boldsymbol{B}(t)\|<K$, $\|\boldsymbol{C}(t)\|<K$, $\forall t\in \boldsymbol{R}$

(ii) 对于某个 σ

$$0<\alpha_0(\sigma)\boldsymbol{I}\leqslant \boldsymbol{W}_C(t,\ t+\sigma)\quad \forall t\in \boldsymbol{R}$$

(iii) 对于某个 σ'

$$0<\alpha_0'(\sigma')\boldsymbol{I}\leqslant \boldsymbol{W}_0(t,\ t+\sigma).\quad \forall t\in \boldsymbol{R}$$

B-I-11 特征值.特征向量

在任意体 F 上，考察矩阵 $\boldsymbol{A}\in F^{n\times n}$ 和列向量 $\boldsymbol{x}\in F^n$. $\boldsymbol{A}\boldsymbol{x}\in F^n$，但一般 $\boldsymbol{A}\boldsymbol{x}$ 是和 $\boldsymbol{x}$ 不同的向量. 现在，我们特别来考察 $\boldsymbol{A}\boldsymbol{x}$ 和向量 $\boldsymbol{x}$ 成比例，即对于某个 $\lambda\in F$，

$$\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x} \tag{1}$$

成立的情况. 若 $\boldsymbol{x}$ 是零向量，当然(1)式成立. 我们不讨论这种情况，即假定 $\boldsymbol{x}\neq\boldsymbol{\theta}$.

当有满足(1)式的向量 $\boldsymbol{x}\neq\boldsymbol{\theta}$ 存在时，λ 称为矩阵 $\boldsymbol{A}$ 的特征值(eigen value)，$\boldsymbol{x}$ 称为对应的(右)特征向量(eigenvector).

现假定对应于某个特征值 λ 有两个线性独立的特征向量 $\boldsymbol{x}_1$, $\boldsymbol{x}_2$，则 $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ 的任意线性组合仍然是对应于 λ 的特征向量

$$\begin{aligned}\boldsymbol{A}(\alpha_1\boldsymbol{x}_1+\alpha_2\boldsymbol{x}_2)&=\alpha_1\boldsymbol{A}\boldsymbol{x}_1+\alpha_2\boldsymbol{A}\boldsymbol{x}_2=\alpha_1\lambda\boldsymbol{x}_1+\alpha_2\lambda\boldsymbol{x}_2\\&=\lambda(\alpha_1\boldsymbol{x}_1+\alpha_2\boldsymbol{x}_2)\quad {}^\forall\alpha_1,\ \alpha_2\in F\end{aligned} \tag{2}$$

虽然满足(1)式的线性独立的向量只有一个，但是特征向量毕竟不是唯一的. 其原因是，若 $\boldsymbol{x}$ 是特征向量，则 $\alpha\boldsymbol{x}\,(\alpha\in F,\ \alpha\neq 0)$ 也是特征向量. 为了唯一确定特征向量，必须加上某些规格化条件.

将(1)式变化后得

$$(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0} \tag{3}$$

由 B-I-4, (3)式具有非零向量解的充分必要条件是

$$\det(\lambda\boldsymbol{I}-\boldsymbol{A})=0. \tag{4}$$

(4)式一般称为 $\boldsymbol{A}$ 的特征方程式，(4)式左边的 $\det(\lambda\boldsymbol{I}-\boldsymbol{A})$ 称为特征多项式. 因此有下列定理.

［定理 1］ λ 是矩阵 $\boldsymbol{A}$ 的特征值的充分必要条件是，λ 是 $\boldsymbol{A}$ 的特征多项式的根.

但是必须注意，特征多项式的根不一定在体 F 内. 例如，即使 F 是实数体 R、$\boldsymbol{A}\in R^{n\times n}$ 的特征方程式的根说不定是复数. 在这种情况下，满足(1)式的 $\boldsymbol{x}\in R^n$ 不存在.

以下为了简单起见，假定 F 为复数体 C，矩阵 $\boldsymbol{A}\in C^{n\times n}$. 当然，作为特殊情况也包括 $\boldsymbol{A}\in R^{n\times n}$. 但是，无论在那一种情况下，关于特征值 λ、特征向量 $\boldsymbol{x}$，一般 $\lambda\in C$, $\boldsymbol{x}\in C^n$. 若令 $F=C$，由代数方程式的性质可知，(4)式的根，即 $\boldsymbol{A}$ 的特征值，包括重根在内，共有 n 个.

矩阵 $\boldsymbol{A}$ 的特征值常写成 $\lambda(\boldsymbol{A})$. 当对 n 个特征值编号加以区别时，第 i 个特征值写成 $\lambda_i(\boldsymbol{A})$. 此外，当特征值均为实数，大小关系确定时，将最大(小)的特征值写成 $\lambda_{\max}(\boldsymbol{A})$ $[\lambda_{\min}(\boldsymbol{A})]$，而且将 $\boldsymbol{A}$ 的 n 个特征值全体的集合写成 $\Lambda(\boldsymbol{A})$.

例如

［系 1］ 对于任意 $\boldsymbol{A}\in C^{n\times n}$，若 λ 是 $\boldsymbol{A}$ 的特征值，则也是 $\boldsymbol{A}^T$ 的特征值. 其逆亦真.

利用上述符号，可将［系 1］表示成下列［系 1′］或［系 1″］.

［系 1′］ 若进行适当编号，则 $\lambda_i(\boldsymbol{A})=\lambda_i(\boldsymbol{A}^T)\ (i=1,\ \cdots,\ n)$, ${}^\forall\boldsymbol{A}\in C^{n\times n}$ 成立.

［系 1″］ $\Lambda(\boldsymbol{A})=\Lambda(\boldsymbol{A}^T)\ {}^\forall\boldsymbol{A}\in C^{n\times n}$.

下面给出根据定义本身及定理 1 很容易推导出来的性质.

[系 2] 若进行适当编号, 则 $\lambda_i(\boldsymbol{A}')=\lambda_i(\boldsymbol{A})$ $(i=1, \cdots, n)$ 成立.

[系 3] $\boldsymbol{A}$ 是非正则的 $\Leftrightarrow$ 至少 $\lambda_i(\boldsymbol{A})$ 的一个是 0.

[系 4] 若适当进行编号, 则 $\lambda_i(\boldsymbol{A}-\alpha\boldsymbol{I})=\lambda_i(\boldsymbol{A})-\alpha$ $(i=1, \cdots, n)$ 成立 $({}^\forall\alpha\in F)$.

[系 5] 若 $\boldsymbol{A}$ 为三角形矩阵, 则 $\Lambda(\boldsymbol{A})=\{a_{ii}; i=1, \cdots, n\}$.

[系 6] 若 $\boldsymbol{A}$ 为分块三角形矩阵

$$\boldsymbol{A}=\begin{bmatrix} \boldsymbol{A}_{11} & \cdots & \boldsymbol{A}_{n1} \\ & \ddots & \vdots \\ \boldsymbol{0} & & \boldsymbol{A}_{nn} \end{bmatrix}$$

则 $\Lambda(\boldsymbol{A})=\bigcup\limits_{i=1}^{n}\Lambda(\boldsymbol{A}_{ii})$.

[系 7] 若适当进行编号, 则 $\lambda_i(\boldsymbol{A}^k)=[\lambda_i(\boldsymbol{A})]^k$, $k\geqslant1$, $i=1, \cdots, n$. 因此, 幂零矩阵的特征值全为 0.

[系 8] 设 $\boldsymbol{A}$ 为正则的. 若适当进行编号, 则 $\lambda_i(\boldsymbol{A}^{-1})=1/[\lambda_i(\boldsymbol{A})]$, $i=1, \cdots, n$.

(注意)

由系 7, 8 立即可见, $[\lambda_i(\boldsymbol{A})]^k$, $1/[\lambda_i(\boldsymbol{A})]$ 分别是 $\boldsymbol{A}^k$, $\boldsymbol{A}^{-1}$ 的特征值. 反过来, $\boldsymbol{A}^k$, $\boldsymbol{A}^{-1}$ 的特征值只有上面形式这一点并不是很明显的. 但是若利用下面将要讲的若当标准形, 可以很简单地证明.

此外, 与 $\boldsymbol{A}^k$ 的特征值 $[\lambda_i(\boldsymbol{A})]^k$ 相对应的特征向量, 是与 $\boldsymbol{A}$ 的 $\lambda_i(\boldsymbol{A})$ 相对应的特征向量. 其逆亦真. 关于 $\boldsymbol{A}^{-1}$ 也同样.

[系 9] 相似矩阵具有相同的特征值, 即 $\Lambda(\boldsymbol{A})=\Lambda(\boldsymbol{TAT}^{-1})$ 对于任意正则矩阵 $\boldsymbol{T}$ 均成立.

[系 10] 对于任意 $\boldsymbol{A}\in F^{m\times n}$, $\boldsymbol{B}\in F^{n\times m}$, 在 $\boldsymbol{AB}\in F^{m\times m}$ 和 $\boldsymbol{BA}^*\in F^{n\times n}$ 之间, 除了 0 以外, 其特征值一致. 即若对除了 0 以外的特征值适当进行编号, 则 $\lambda_i(\boldsymbol{AB})=\lambda_i(\boldsymbol{BA})$. 因此, 若 $m<n[m>n]$, 则 $\Lambda(\boldsymbol{AB})$ $[\Lambda(\boldsymbol{BA})]$ 至少包含有 $n-m$ 个 $[m-n$ 个$]$ 0.

设 $\Lambda(\boldsymbol{A})=\{\lambda_i; i=1, \cdots, n\}$, 与 λ_i 对应的特征向量为 $\boldsymbol{x}_i$. 对于 $\boldsymbol{B}\in F^{m\times m}$ 也同样, 设 $\Lambda(\boldsymbol{B})=\{\mu_j; j=1, \cdots, m\}$, 与 μ_j 对应的特征向量为 $\boldsymbol{y}_i$. 这时有下列性质.

[系 11] $\boldsymbol{x}_i\otimes\boldsymbol{y}_j$ $(i=1, \cdots, n; j=1, \cdots, m)$ 是 $\boldsymbol{A}\otimes\boldsymbol{B}$ 的特征向量, $\lambda_i\mu_j$ 是与其对应的特征值.

[系 12] $\boldsymbol{x}_i\otimes\boldsymbol{y}_j$ $(i=1, \cdots, n; j=1, \cdots, m)$ 是 $(\boldsymbol{A}\otimes\boldsymbol{I}_m)+(\boldsymbol{I}_n\otimes\boldsymbol{B})$ 的特征向量, $\lambda_i+\mu_j$ 是与其对应的特征值.

(问题 1) 试证明 [系 1]—[系 12].

(问题 2) 试证明实正交矩阵的特征值限于 ±1.

(注意)

(i) 将系 3 和系 11 合起来可以看到, 克罗内克积是正则的充分必要条件是, $\boldsymbol{A}$, $\boldsymbol{B}$ 都是正则的 (参看 B-I-4 中克罗内克积).

(ii) 根据系 6 和系 9, 若 $\boldsymbol{A}$ 相似于分块三角形矩阵, 则可不必去求大矩阵的特征值, 而来求一些小矩阵的特征值, 这样将使计算量大大减少. 作为特殊情况, 当 $\boldsymbol{T}$ 是顺列矩

* 原文误印成 $\boldsymbol{B}\in \boldsymbol{A}F^{n\times n}$. ——译者注

阵时，即适当地交换行和列的顺序，可以变成分块三角形矩阵的情况下，该性质特别有效。可以利用和 A-I-4 中讲过的相同方法，确定使 $\tilde{\boldsymbol{A}}=\boldsymbol{P}^T\boldsymbol{A}\boldsymbol{P}$ 变成分块三角形的顺列矩阵 $\boldsymbol{P}$ 是否存在.

$\Lambda(\boldsymbol{A})$ 有时重复地包含同一个数，即为特征方程式含有重根的情况．在 n 个特征值中，假定相异的有 σ 个，将其写成 $\lambda_j(j=1,\cdots,\sigma)$[1]．当 $\Lambda(\boldsymbol{A})$ 中包含有 m_j 个 λ_j，即 λ_j 是特征方程式的 m_j 重根时，λ_j 作为 $\boldsymbol{A}$ 的特征值，称为具有代数重复度 m_j．对此有下列定理.

［定理 2］　对于 $\boldsymbol{A}\in F^{n\times n}$ 的特征值 λ_j，设 $\operatorname{rank}[\lambda_j\boldsymbol{I}-\boldsymbol{A}]=n-\alpha_j$，则有 $m_j\geqslant\alpha_j$.

因证明太长，这里省略(例如可参看文献[60])．上述的 α_j 叫做 λ_j 的几何重复度．对于所有相异的特征值，当 $m_j=\alpha_j$，即代数重复度和几何重复度相等时，称为该矩阵是单纯矩阵(simple matrix).

(例)

$$\begin{bmatrix}\lambda_1&1&0&0\\0&\lambda_1&0&0\\0&0&\lambda_2&1\\0&0&0&\lambda_2\end{bmatrix}\begin{bmatrix}\lambda_1&0&0&0\\0&\lambda_1&0&0\\0&0&\lambda_2&1\\0&0&0&\lambda_2\end{bmatrix}\begin{bmatrix}\lambda_1&1&0&0\\0&\lambda_1&0&0\\0&0&\lambda_2&0\\0&0&0&\lambda_2\end{bmatrix}\begin{bmatrix}\lambda_1&0&0&0\\0&\lambda_1&0&0\\0&0&\lambda_2&0\\0&0&0&\lambda_2\end{bmatrix}.$$

其中每个矩阵都是 $\sigma=2$，$m_1=m_2=2$．按从左到右的顺序，$\alpha_1=\alpha_2=1$；$\alpha_1=2$，$\alpha_2=1$；$\alpha_1=1$，$\alpha_2=2$；$\alpha_1=\alpha_2=2$．即仅仅最右边的矩阵是单纯的.

若所有特征值均相异，该矩阵当然是单纯矩阵(为何？)

(问题 3)　试证明

$$\boldsymbol{A}\text{ 是单纯矩阵}\Leftrightarrow\boldsymbol{A}^T\text{ 是单纯矩阵}$$

［引理 1］　设 λ_j 的几何重复度为 α_j，则对应于 λ_j 有 α_j 个线性独立的特征向量存在，$\boldsymbol{A}$ 的任意特征向量均可用其线性组合表示.

(证明)　根据(3)式和 B-I-4 中定理 3.

［定理 3］　设 $\boldsymbol{A}$ 相异的特征值为 $\lambda_1,\cdots,\lambda_\sigma$，与 λ_j 对应线性独立的特征向量为 $\boldsymbol{x}_1^j,\cdots,\boldsymbol{x}_{\alpha_j}^j$，则 $\boldsymbol{x}_1^1,\cdots,\boldsymbol{x}_{\alpha_1}^1,\boldsymbol{x}_1^2,\cdots,\boldsymbol{x}_1^\sigma,\cdots,\boldsymbol{x}_{\alpha_\sigma}^\sigma$ 是线性独立的.

(证明)　对于某组数 $C_{jk}\in F(k=1,\cdots,\alpha_j;\ j=1,\cdots,\sigma)$，设

$$C_{11}\boldsymbol{x}_1^1+\cdots+C_{1\alpha_1}\boldsymbol{x}_{\alpha_1}^1+C_{21}\boldsymbol{x}_1^2+\cdots+C_{\sigma1}^{\sigma}\boldsymbol{x}_{\sigma1}+\cdots+C_{\sigma\alpha_\sigma}\boldsymbol{x}_{\alpha\sigma}^\sigma=\boldsymbol{0}$$

成立．对上式左乘以 $(\lambda_1\boldsymbol{I}-\boldsymbol{A})$，利用 $(\lambda_1\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}_k^1=\boldsymbol{0}(k=1,\cdots,\alpha_1)$、$\boldsymbol{A}\boldsymbol{x}_k^j=\lambda_j\boldsymbol{x}_k^j(k=1,\cdots,\alpha_i;\ j=2,\cdots,s;)$，则

$$(\lambda_1-\lambda_2)C_{21}\boldsymbol{x}_1^2+\cdots+(\lambda_1-\lambda_\sigma)C_{\sigma\alpha_\sigma}\boldsymbol{x}_{\alpha\sigma}^\sigma=\boldsymbol{0}.$$

以下依次左乘以 $(\lambda_2\boldsymbol{I}-\boldsymbol{A}),\cdots,(\lambda_{\sigma-1}\boldsymbol{I}-\boldsymbol{A})$．最后，因

---

[1] 特征值有两种编号方法．一种是，不管有无重根，都按着 $\lambda_1,\cdots,\lambda_n$ 编；另一种是，在有重根的情况下，对相同的特征值给以同样的编号，即 $\lambda_1,\cdots,\lambda_\sigma$．例如，对于

$$\boldsymbol{A}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&2\end{bmatrix}.$$

按前一种方法是，$\lambda_1=1$，$\lambda_2=1$，$\lambda_3=2$．若按后一种方法，则为 $\lambda_1=1$(2 重根)，$\lambda_2=2$．为了区别起见，将前者写为“特征值 $\lambda_i(i=1,\cdots,n)$”，将后者写为“相异的特征值 $\lambda_j(j=1,\cdots,\sigma)$”．特别是，在容易混淆的情况下，要将前者写成 $\lambda_{(1)},\cdots,\lambda_{(n)}$.

$$(\lambda_1-\lambda_\sigma)(\lambda_2-\lambda_\sigma)\cdots(\lambda_{\sigma-1}-\lambda_\sigma)(C_{\sigma1}\boldsymbol{x}_1^\sigma+\cdots+C_{\sigma\alpha_\sigma}\boldsymbol{x}_{\alpha\sigma}^\sigma)=\mathbf{0}$$

$$\lambda_j\neq\lambda_\sigma,\ j=1,\ \cdots,\ \sigma-1$$

则

$$C_{\sigma1}\boldsymbol{x}_1^\sigma+\cdots+C_{\sigma\alpha_\sigma}\boldsymbol{x}_{\alpha_\sigma}^\sigma=\mathbf{0}$$

因为 $\boldsymbol{x}_1^\sigma,\cdots,\boldsymbol{x}_{\alpha}^\sigma$ 是线性独立的，则 $C_{\sigma1}=\cdots=C_{\sigma\alpha_\sigma}=0$. 重复同样操作，最后可以证明所有的 C_{jk} 都必须是 0.

证明完毕

[系 13] 与 $\boldsymbol{A}$ 的相异特征值对应的特征向量是线性独立的.

[系 14] $\boldsymbol{A}$ 最多有 $\left(\sum_{j=1}^{\sigma}\alpha_j\right)$ 个线性独立的特征向量存在.

[系 15] $\boldsymbol{A}$ 是单纯矩阵 $\Leftrightarrow\boldsymbol{A}$ 有 n 个线性独立的特征向量存在.

设 $\boldsymbol{A}$ 有 n 个线性独立的特征向量 $\boldsymbol{x}_1,\ \cdots,\ \boldsymbol{x}_n$, 以这些向量为列组成矩阵 $\boldsymbol{T}$

$$\boldsymbol{T}\triangleq[\boldsymbol{x}_1,\ \boldsymbol{x}_2,\ \cdots,\ \boldsymbol{x}_n] \tag{5}$$

则

$$\boldsymbol{AT}=[\boldsymbol{Ax}_1,\ \cdots,\ \boldsymbol{Ax}_n]=[\lambda_1\boldsymbol{x}_1,\ \cdots,\ \lambda_n\boldsymbol{x}_n]$$

$$=[\boldsymbol{x}_1,\ \cdots,\ \boldsymbol{x}_n]\begin{bmatrix}\lambda_1 & & \mathbf{0}\\ & \ddots & \\ \mathbf{0} & & \lambda_n\end{bmatrix}=\boldsymbol{T\Lambda} \tag{6}$$

成立. 式中 λ_i 是对应于 $\boldsymbol{x}_i$ 的特征值, Λ 是以 λ_i 为对角线元素的对角矩阵. 上式两边右乘以 $\boldsymbol{T}^{-1}$, 得

$$\boldsymbol{A}=\boldsymbol{T\Lambda T}^{-1} \tag{7}$$

由该结果和系 15 得出下列重要结果.

[定理 4] $\boldsymbol{A}$ 是单纯矩阵 $\Leftrightarrow\boldsymbol{A}$ 相似于对角形矩阵.

(证明) $\Rightarrow$上面已证过, $\Leftarrow$可按与上面证明相反方向进行.

如问题 2 所述，若 $\boldsymbol{A}$ 是单纯矩阵，则 $\boldsymbol{A}^T$ 也是单纯矩阵. 因此，由系 14, $\boldsymbol{A}^T$ 也有 n 个线性独立的特征向量 $\boldsymbol{y}_1,\ \cdots,\ \boldsymbol{y}_n$ 存在. 利用系 1, 得

$$\boldsymbol{A}^T\boldsymbol{y}_i=\lambda_i\boldsymbol{y}_i\quad i=1,\ \cdots,\ n \tag{8}$$

将(8)式两边取转置，得

$$\boldsymbol{y}_i^T\boldsymbol{A}=\lambda_i\boldsymbol{y}_i^T \tag{9}$$

因此, $\boldsymbol{y}_i^T$ 称为 $\boldsymbol{A}$ 的左特征向量. 根据系 1, 无需对特征值的左右加以区别.

作(5)式 $\boldsymbol{T}$ 的逆矩阵 $\boldsymbol{T}^{-1}$, 设其第 i 行为 $\boldsymbol{y}_i^T$, 即

$$\boldsymbol{T}^{-1}=\begin{bmatrix}\boldsymbol{y}_1^T\\ \vdots\\ \boldsymbol{y}_n^T\end{bmatrix}.$$

现在我们来证明, $\boldsymbol{y}_i^T$ 就是 $\boldsymbol{A}$ 的左特征向量. (7)式左乘以 $\boldsymbol{y}_i^T$, 得

$$\boldsymbol{y}_i^T\boldsymbol{A}=\boldsymbol{y}_i^T\boldsymbol{T\Lambda T}^{-1}=\boldsymbol{e}_i^T\boldsymbol{\Lambda T}^{-1}=\lambda_i\boldsymbol{e}_i^T\boldsymbol{T}^{-1}=\lambda_i\boldsymbol{y}_i^T\quad(i=1,\ \cdots,\ n)$$

因此 $\boldsymbol{y}_i^T$ 是 $\boldsymbol{A}$ 的左特征向量.

再由(7)式

$$\boldsymbol{A}=\boldsymbol{T}\boldsymbol{\Lambda}\boldsymbol{T}^{-1}=[\boldsymbol{x}_1\cdots\boldsymbol{x}_n]\begin{bmatrix}\lambda_1 & & \boldsymbol{0}\\ & \ddots & \\ \boldsymbol{0} & & \lambda_n\end{bmatrix}\begin{bmatrix}\boldsymbol{y}_1^T\\ \vdots\\ \boldsymbol{y}_n^T\end{bmatrix}=[\lambda_1\boldsymbol{x}_1\cdots\lambda_n\boldsymbol{x}_n]\begin{bmatrix}\boldsymbol{y}_1^T\\ \vdots\\ \boldsymbol{y}_n^T\end{bmatrix} \tag{10}$$

利用 B-I-1 中分块矩阵的乘法运算 d，得

$$\boldsymbol{A}=\sum_{i=1}^{n}\lambda_i\boldsymbol{G}_i \quad \boldsymbol{G}_i\triangleq\boldsymbol{x}_i\boldsymbol{y}_i^T=n\times n \quad i=1,\ \cdots,\ n \tag{11}$$

它称为单纯矩阵 $\boldsymbol{A}$ 的谱分解.

(11)式的右边，与特征值相同或相异毫无关系，总之是在假定有 n 个线性独立的特征向量存在下得出来的. 现将上述内容再稍微整理一下. 和定理 3 一样，假定 $\boldsymbol{A}$ 有 σ 个相异的特征值 $\lambda_1,\ \cdots,\ \lambda_\sigma$，对应于 λ_j 线性独立的右特征向量为 $\boldsymbol{x}_1^j,\ \cdots,\ \boldsymbol{x}_{\alpha j}^j$，同样左特征向量为 $\boldsymbol{y}_1^{jT},\ \cdots,\ \boldsymbol{y}_{\alpha j}^{jT}$，则(11)式可写成

$$\boldsymbol{A}=\sum_{j=1}^{\sigma}\lambda_j\sum_{k=1}^{\alpha j}\boldsymbol{x}_k^j\boldsymbol{y}_k^{jT} \tag{12}$$

在这里，若定义

$$\boldsymbol{E}_j\triangleq\sum_{k=1}^{\alpha_j}\boldsymbol{x}_k^j\boldsymbol{y}_k^{jT}=[\boldsymbol{x}_1^j\ \cdots\ \boldsymbol{x}_{\alpha_j}^j]\begin{bmatrix}\boldsymbol{y}_1^{jT}\\ \vdots\\ \boldsymbol{y}_{\alpha j}^{jT}\end{bmatrix} \tag{13}$$

则下列定理成立.

［定理 5］ 对于单纯矩阵，能够进行满足下列性质的谱分解.

(i) $\boldsymbol{A}=\sum_{j=1}^{\sigma}\lambda_j\boldsymbol{E}_j$ (14)

(ii) $\boldsymbol{E}_j^2=\boldsymbol{E}_j \quad j=1,\ \cdots,\ \sigma$ (15)

(iii) $\boldsymbol{E}_j\boldsymbol{E}_k=\boldsymbol{0} \quad j,\ k=1,\ \cdots,\ \sigma;\ j\neq k$ (16)

(iv) $\sum_{j=1}^{\sigma}\boldsymbol{E}_j=\boldsymbol{I}$ (17)

(v) 满足性质(i)～(iv)的矩阵 $\boldsymbol{E}_j(j=1,\ \cdots,\ \sigma)$ 是唯一确定的.

(vi) rank $\boldsymbol{E}_j=\alpha_j \quad j=1,\ \cdots,\ \sigma$ (18)

（证明） (i)上面已证过. 根据 $\boldsymbol{y}_k^{jT}$($\boldsymbol{T}^{-1}$ 的行)和 $\boldsymbol{x}_l^{j'}$($\boldsymbol{T}$ 的列)的正交性

$$\boldsymbol{y}_k^{jT}\boldsymbol{x}_l^{j'}=\delta_{jk}\delta_{j'l}$$
$$k=1,\ \cdots,\ \alpha_j;\ l=1,\ \cdots,\ \alpha_{j'};\ j,\ j'=1,\ \cdots,\ \sigma$$

立即可以得出 $\boldsymbol{E}_j$ 的性质(ii)，(iii)，(iv).

$$\boldsymbol{I}=\boldsymbol{T}\boldsymbol{T}^{-1}=[\boldsymbol{x}_1^1\cdots\boldsymbol{x}_{j\alpha_\sigma}^\sigma]\begin{bmatrix}\boldsymbol{y}_1^{1T}\\ \vdots\\ \boldsymbol{y}_{\alpha\sigma}^{\sigma T}\end{bmatrix}=\sum_{j=1}^{\sigma}\sum_{k=1}^{\alpha_\sigma}\boldsymbol{x}_k^j\boldsymbol{y}_k^{jT}=\sum_{j=1}^{\sigma}\boldsymbol{E}_j$$

假定对于和 $\boldsymbol{E}_j$ 不同的矩阵 $\widetilde{\boldsymbol{E}}_j$ 性质(i)～(iv)也成立. 由(i)，(ii)，(iii)，得

$$\boldsymbol{E}_j\boldsymbol{A}=\boldsymbol{A}\boldsymbol{E}_j=\lambda_j\boldsymbol{E}_j,\ \widetilde{\boldsymbol{E}}_{j'}\boldsymbol{A}=\boldsymbol{A}\widetilde{\boldsymbol{E}}_{j'}=\lambda_{j'}\widetilde{\boldsymbol{E}}_{j'}$$

$$\therefore\quad \boldsymbol{E}_j(\boldsymbol{A}\widetilde{\boldsymbol{E}}_{j'})=\lambda_{j'}\boldsymbol{E}_j\widetilde{\boldsymbol{E}}_{j'}{}^{*}$$

$$(\boldsymbol{E}_j\boldsymbol{A})\widetilde{\boldsymbol{E}}_{j'}=\lambda_j\boldsymbol{E}_j\widetilde{\boldsymbol{E}}_{j'}{}^{**}$$

$$\therefore\quad (\lambda_{j'}-\lambda_j)\boldsymbol{E}_j\widetilde{\boldsymbol{E}}_{j'}=\boldsymbol{0}$$

* 原文误印成 $\boldsymbol{E}_j(\boldsymbol{A}\widetilde{\boldsymbol{E}}_{j'})=\lambda_j\widetilde{\boldsymbol{E}}_j\widetilde{\boldsymbol{E}}_{j'}$. ——译者注

** 原文误印成 $(\boldsymbol{E}_j\boldsymbol{A})\boldsymbol{E}_{j'}=\lambda_j\boldsymbol{E}_j\boldsymbol{E}_{j'}$. ——译者注

因 $\lambda_{j'}\neq\lambda_j$，则 $\boldsymbol{E}_j\widetilde{\boldsymbol{E}}_{j'}=\mathbf{0}$. 由此得

$$\boldsymbol{E}_j=\boldsymbol{E}_j\boldsymbol{I}=\boldsymbol{E}_j\sum_{j'=1}^{\sigma}\widetilde{\boldsymbol{E}}_{j'}=\boldsymbol{E}_j\widetilde{\boldsymbol{E}}_j=\sum_{j'=1}^{\sigma}\boldsymbol{E}_{j'}\widetilde{\boldsymbol{E}}_j=\boldsymbol{I}\widetilde{\boldsymbol{E}}_j=\widetilde{\boldsymbol{E}}_j$$

因此，(v) 成立.

由 (13) 式可见，$\boldsymbol{E}_j$ 是 $n\times\alpha_j$ 矩阵和 $\alpha_j\times n$ 矩阵之积. 因此，由 B–I–3 中 (23) 式，得

$$\operatorname{rank}\boldsymbol{E}_j\leqslant\alpha_j,\quad j=1,\ \cdots,\ \sigma$$

此外，由 (iv) 及 B–I–3 中 (18) 式，得

$$n=\sum_{j=1}^{\sigma}\alpha_j=\operatorname{rank}\boldsymbol{I}\leqslant\sum_{j=1}^{\sigma}\operatorname{rank}\boldsymbol{E}_j$$

由上两式，则 (vi) 成立.

证明完毕

我们知道，单纯矩阵与对角矩阵这样一种简单形状矩阵相似，那么非单纯矩阵是否也能和一个简单形状的矩阵相似呢？下面的定理回答了这个问题.

［定理 6］ 任意 $\boldsymbol{A}\in C^{n\times n}$ 和下列分块对角矩阵相似：

$$\boldsymbol{J}=\begin{bmatrix}\boldsymbol{J}_1 & & & \mathbf{0}\\ & \boldsymbol{J}_2 & & \\ & & \ddots & \\ \mathbf{0} & & & \boldsymbol{J}_\sigma\end{bmatrix}=n\times n \tag{19}$$

$$\boldsymbol{J}_j=\begin{bmatrix}\boldsymbol{J}_{j1} & & & \mathbf{0}\\ & \boldsymbol{J}_{j2} & & \\ & & \ddots & \\ \mathbf{0} & & & \boldsymbol{J}_{j\alpha_j}\end{bmatrix}=m_j\times m_j\quad j=1,\ \cdots,\ \sigma \tag{20}$$

$$\boldsymbol{J}_{jk}=\begin{bmatrix}\lambda_j & 1 & & \mathbf{0}\\ & \ddots & \ddots & \\ & & \ddots & 1\\ \mathbf{0} & & & \lambda_j\end{bmatrix}=n_{jk}\times n_{jk}\quad \begin{matrix}k=1,\ \cdots,\ \alpha_j\\ j=1,\ \cdots,\ \sigma,\end{matrix} \tag{21}$$

在这里，$\boldsymbol{J}_{jk}=n_{jk}\times n_{jk}$；如上所示，其主对角线上的元素是第 j 个相异的特征值 λ_j，主对角线上面一条对角线上的元素全为 1，除此之外，$\boldsymbol{J}_{jk}$ 中的其它元素全为 0. $\boldsymbol{J}_j$：该典型矩阵是一个主对角线上排列着子块的对角矩阵，其子块个数等于 λ_j 的几何重复度 α_j. 而且，$\boldsymbol{J}_j$ 的行和列数，即 $\sum\limits_{k=1}^{\alpha_j}n_{jk}$ 等于 λ_j 的代数重复度 m_j. $\boldsymbol{J}$ 是以 $\boldsymbol{J}_j$ 为主对角线上子块的分块对角矩阵，其子块的个数等于 $\boldsymbol{A}$ 的相异特征值的数目 σ. 而且，$\boldsymbol{J}$ 的行和列数，即 $\sum\limits_{j=1}^{\sigma}m_j=\sum\limits_{j=1}^{\sigma}\sum\limits_{k=1}^{\alpha_j}n_{jk}$ 当然等于 n. 该矩阵 $\boldsymbol{J}$ 称为若当标准形 (Jorda's canonical form).

显然，若当标准形包含对应于单纯矩阵的对角形，后者是若当标准形的特殊情况.

定理 6 的证明可以参看文献 [10, 25]，这里省略. 根据该定理，则对于任意 $\boldsymbol{A}\in C^{n\times n}$，有使

$$\boldsymbol{AT}=\boldsymbol{TJ} \tag{22}$$

的 $\boldsymbol{T}\in C^{n\times n}$, $\det\boldsymbol{T}\neq 0$ 存在. 仿照 $\boldsymbol{J}$ 对式中 $\boldsymbol{T}$ 分割, 分成如下分块:

$$\boldsymbol{T}=[\boldsymbol{T}_1 \vdots \cdots \vdots \boldsymbol{T}_\sigma] \quad (\boldsymbol{T}_j=n\times m_j) \tag{23}$$

$$\boldsymbol{T}_j=[\boldsymbol{T}_{j1} \vdots \cdots\cdots \vdots \boldsymbol{T}_{j\alpha_j}] \quad (\boldsymbol{T}_{jk}=n\times n_{jk}) \tag{24}$$

将其代入(22)式, 考虑到 $\boldsymbol{J}$ 的组成, 可以看出

$$\boldsymbol{A}\boldsymbol{T}_{jk}=\boldsymbol{T}_{jk}\boldsymbol{J}_{jk} \quad k=1,\ \cdots,\ \alpha_j;\ j=1,\ \cdots,\ \sigma \tag{25}$$

设 $\boldsymbol{T}_{jk}$ 的第 1 列为 $\boldsymbol{t}_{jk}^1$, 由(21), (25)式得

$$\boldsymbol{A}\boldsymbol{t}_{jk}^1=\lambda_j\boldsymbol{t}_{jk}^1. \tag{26}$$

即 $\boldsymbol{t}_{jk}^1$ 是对应于 λ_j 的 $\boldsymbol{A}$ 的特征向量之一. 其次, 对于 $\boldsymbol{T}_{jk}$ 的第 2 列 $\boldsymbol{t}_{jk}^2$, 仍由 (21), (25) 式, 得

$$\boldsymbol{A}\boldsymbol{t}_{jk}^2=\boldsymbol{t}_{jk}^1+\lambda_j\boldsymbol{t}_{jk}^2 \tag{27}$$

于是

$$(\boldsymbol{A}-\lambda_j\boldsymbol{I})\boldsymbol{t}_{jk}^2=\boldsymbol{t}_{jk}^1\neq\boldsymbol{0} \tag{28}$$

因此, $\boldsymbol{t}_{jk}^2$ 不是特征向量, 但是

$$(\boldsymbol{A}-\lambda_j\boldsymbol{I})^2\boldsymbol{t}_{jk}^2=(\boldsymbol{A}-\lambda_j\boldsymbol{I})\boldsymbol{t}_{jk}^1=\boldsymbol{0} \tag{29}$$

也就是说, $\boldsymbol{t}_{jk}^2$ 是这样一种向量, 虽然它乘一次 $(\boldsymbol{A}-\lambda_j\boldsymbol{I})$ 不为 $\boldsymbol{0}$, 但乘 2 次即变成 $\boldsymbol{0}$. 对于以下的 $\boldsymbol{t}_{jk}^3$, $\boldsymbol{t}_{jk}^4$ 等等, 也可以用同样的方法证明. 一般, 虽然 $\boldsymbol{t}_{jk}^l$ 乘上 $(\boldsymbol{A}-\lambda_j\boldsymbol{I})^{l-1}$ 不为 $\boldsymbol{0}$, 但乘上 $(\boldsymbol{A}-\lambda_j\boldsymbol{I})^l$ 即变成 $\boldsymbol{0}$, 而且有下列关系

$$(\boldsymbol{A}-\lambda_j\boldsymbol{I})\boldsymbol{t}_{jk}^l=\boldsymbol{t}_{jk}^{(l-1)} \tag{30}$$

$\boldsymbol{t}_{jk}^l(l\geqslant 2)$ 通常称为广义特征向量.

利用这些关系可以作出矩阵 $\boldsymbol{T}$, 而且还有稍简单的方法[61]. 实际上, 并非一定要把所给矩阵 $\boldsymbol{A}$ 变换成若当形, 但是和若当形相似是很重要的, 在这个基础上可以得出许多重要的结论.

(问题 4) 利用和若当形相似, 试证明

$$\det\boldsymbol{A}=\prod_{i=1}^{n}\lambda_i=\prod_{j=1}^{\sigma}(\lambda_j)^{m_j}$$

$$\mathrm{tr}\,\boldsymbol{A}=\sum_{i=1}^{n}\lambda_i=\sum_{j=1}^{\sigma}m_j\lambda_j.$$

设矩阵 $\boldsymbol{A}$ 的若当形为 $\boldsymbol{J}$, 由(22)式得

$$\boldsymbol{A}=\boldsymbol{T}\boldsymbol{J}\boldsymbol{T}^{-1} \tag{31}$$

而且

$$\boldsymbol{A}^p=\boldsymbol{T}\boldsymbol{J}^p\boldsymbol{T}^{-1} \quad p=1,\ 2,\ \cdots \tag{32}$$

$$\boldsymbol{A}^{-1}=\boldsymbol{T}\boldsymbol{J}^{-1}\boldsymbol{T}^{-1} \tag{33}$$

$$(\lambda\boldsymbol{I}-\boldsymbol{A})^p=\boldsymbol{T}(\lambda\boldsymbol{I}-\boldsymbol{J})^p\boldsymbol{T}^{-1} \quad p=1,\ 2,\ \cdots,\ \forall\lambda\in C \tag{34}$$

成立. 上面的证明都是很容易的.

彼此相似的矩阵的秩相等(B-I-3 中系 1). 因此, $(\lambda\boldsymbol{I}-\boldsymbol{A})^p$ 和 $(\lambda\boldsymbol{I}-\boldsymbol{J})^p$ 等秩. 利用该性质, 通过研究 $(\lambda_l\boldsymbol{I}-\boldsymbol{J})^p$ 的秩, 我们来确定 $(\lambda_l\boldsymbol{I}-\boldsymbol{A})^p$ 的秩的性质, 其中 $\lambda_l(1\leqslant l\leqslant\sigma)$ 是 $\boldsymbol{A}$ 的相异特征值之一. 首先, 由(19)—(21)式可见, $(\lambda_l\boldsymbol{I}-\boldsymbol{J})^p$ 是以 $(\lambda_l\boldsymbol{I}-\boldsymbol{J}_{jk})^p$ 为主对角线上元素的分块对角矩阵. 根据分块对角矩阵秩的性质(参看 B-I-3 中(21), (22)式), 则

$$\operatorname{rank}(\lambda_l \boldsymbol{I}-\boldsymbol{J})^p=\sum_{j=1}^{\sigma}\sum_{k=1}^{\alpha_j}\operatorname{rank}(\lambda_l \boldsymbol{I}-\boldsymbol{J}_{jk})^p \quad p=1,\ 2,\ \cdots \tag{35}$$

成立. 由(21)式,因

$$(\lambda_l \boldsymbol{I}-\boldsymbol{J}_{jk})=\begin{bmatrix} (\lambda_l-\lambda_j) & -1 & & \\ & & \mathbf{0} & \\ & & & \\ & \mathbf{0} & & -1 \\ & & & (\lambda_l-\lambda_j) \end{bmatrix} \tag{36}$$

则$(\lambda_l \boldsymbol{I}-\boldsymbol{J}_{jk})$是三角形矩阵. 若$l\neq j$, 即$\lambda_l\neq\lambda_j$, 则主对角线上的元素不为0. 因此, 根据P-I-1中讲过的行列式的性质(viii), 则$\det(\lambda_l \boldsymbol{I}-\boldsymbol{J}_{jk})\neq 0$, 即

$$\operatorname{rank}(\lambda_l \boldsymbol{I}-\boldsymbol{J}_{jk})=n_{jk} \quad l\neq j \tag{37}$$

其次,在$l=j$, 即$\lambda_l=\lambda_j$的情况下,主对角线上的元素全为0,

$$(\lambda_l \boldsymbol{I}-\boldsymbol{J}_{lk})=-\begin{bmatrix} \boldsymbol{e}_2^T \\ \boldsymbol{e}_3^T \\ \vdots \\ \boldsymbol{e}_{n_{lk}-1}^T \\ \mathbf{0} \end{bmatrix}. \tag{38}$$

右边矩阵成为行标准形. 显然,其秩为$n_{lk}-1$, 即

$$\operatorname{rank}(\lambda_l \boldsymbol{I}-\boldsymbol{J}_{lk})=n_{lk}-1 \tag{39}$$

由(35), (37), (38)式,得

$$\operatorname{rank}(\lambda_l \boldsymbol{I}-\boldsymbol{J})=\sum_{j=1}^{\sigma}\sum_{k=1}^{\alpha_j}n_{jk}-\alpha_l=n-\alpha_l \tag{40}$$

与α_l的定义一致.

其次, 我们来看$(\lambda_l \boldsymbol{I}-\boldsymbol{J})^p(p\geqslant 2)$. 由(36)式容易看到, $(\lambda_l \boldsymbol{I}-\boldsymbol{J}_{jk})^p$主对角线上的元素是$(\lambda_l-\lambda_j)^p$, 若$j\neq l$, 即$\lambda_j\neq\lambda_l$, 则不为0, 因此, 和(37)式同样,

$$\operatorname{rank}(\lambda_l \boldsymbol{I}-\boldsymbol{J}_{jk})^p=n_{jk} \quad j\neq l \tag{41}$$

另一方面,由(38)式可见,

$$(\lambda_l \boldsymbol{I}-\boldsymbol{J}_{lk})^p=\begin{cases} (-1)^p\begin{bmatrix} \boldsymbol{e}_{p+1} \\ \vdots \\ \boldsymbol{e}_{n_{lk}-1} \\ \mathbf{0} \end{bmatrix}\}\ p\text{ 个零行} & 1\leqslant p\leqslant n_{lk} \\ \mathbf{0} & n_{lk}\leqslant p \end{cases} \tag{42}$$

因此,和前面同样,显然

$$\operatorname{rank}(\lambda_l \boldsymbol{I}-\boldsymbol{J}_{lk})^p=\begin{cases} n_{lk}-p & 1\leqslant p\leqslant n_{lk} \\ 0 & n_{lk}\leqslant p. \end{cases} \tag{43}$$

将(41), (43)式代入(35)式,便可求出$\operatorname{rank}(\lambda_l \boldsymbol{I}-\boldsymbol{J})^p$.

由以上结果可见, $(\lambda_l \boldsymbol{I}-\boldsymbol{J})^p$的秩和$\boldsymbol{A}$的若当形$\boldsymbol{J}$的结构之间有下列关系.

(i) 根据定义, λ_l的几何重复度α_l可以用下式给出:

$$\alpha_l=n-\operatorname{rank}(\lambda_l \boldsymbol{I}-\boldsymbol{A}) \tag{44}$$

（ii）若定义

$$\beta_l^p \triangleq \operatorname{rank}(\lambda_l \boldsymbol{I}-\boldsymbol{A})^p-\operatorname{rank}(\lambda_l \boldsymbol{I}-\boldsymbol{A})^{p+1}. \tag{45}$$

则 β_l^p 等于 $\boldsymbol{J}_{lk}(k=1,\ \cdots,\ \alpha_l)$ 中满足 $n_{lk} \geqslant p+1$ 的个数.

（iii）若 $\beta_l^{p_0}=0$ 且 $\beta_l^p \neq 0(p=1,\ 2,\ \cdots,\ p_0-1)$，则

$$p_0=\max_k n_{lk} \tag{46}$$

（iv）$\operatorname{rank}(\lambda_l \boldsymbol{I}-\boldsymbol{A})^{p_0}=n-\sum_{k=1}^{\alpha_l} n_{lk}=n-m_l$

式中 m_l 是 λ_l 的代数重复度.

通常根据这些性质可以知道，将 $\boldsymbol{A}$ 变换成若当形时，将会得到什么样的 $\boldsymbol{J}$.

埃尔米特矩阵．实对称矩阵

下面稍介绍一下埃尔米特矩阵，实对称矩阵的特征值，特征向量.

［定理 7］ 埃尔米特矩阵的特征值是实数.

（证明） 设 $\boldsymbol{A} \in C^{n \times n}$ 为埃尔米特矩阵，λ 是它的一个特征值，则有满足

$$\boldsymbol{A}\boldsymbol{x}=\lambda \boldsymbol{x} \tag{47}$$

的 $\boldsymbol{x} \neq \boldsymbol{\theta}$ 存在．取上式两边的共轭转置，因 $\boldsymbol{A}^*=\boldsymbol{A}$，则

$$\boldsymbol{x}^* A=\bar{\lambda} \boldsymbol{x}^* \tag{48}$$

将(47)* 式左乘以 $\boldsymbol{x}^*$，(48)式右乘以 $\boldsymbol{x}$，相减得

$$\lambda \boldsymbol{x}^* \boldsymbol{x}=\bar{\lambda} \boldsymbol{x}^* \boldsymbol{x} \tag{49}$$

因 $\boldsymbol{x} \neq \boldsymbol{\theta}$，故 $\boldsymbol{x}^* \boldsymbol{x}>0$．因此，$\lambda=\bar{\lambda}$.

证明完毕

［系 16］ 实对称矩阵的特征值是实数.

［系 17］ 若 $\boldsymbol{x}$ 是埃尔米特(实对称)矩阵 $\boldsymbol{A}$ 的右特征向量，则 $\boldsymbol{x}^*(\boldsymbol{x}^T)$ 是 $\boldsymbol{A}$ 的左特征向量.

在可以定义内积的向量空间内，当两个向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 的内积等于 0 时，称为 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 正交(orthogonal)．例如，当 $\boldsymbol{x},\ \boldsymbol{y} \in R^n$，$\boldsymbol{x}^T \boldsymbol{y}=0$ 及 $\boldsymbol{x},\ \boldsymbol{y} \in C^n$，$\boldsymbol{x}^* \boldsymbol{y}=0$ 成立时，称为 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 正交.

［定理 8］ 对应于埃尔米特矩阵相异特征值的特征向量彼此正交.

（证明） 设 $\lambda_1,\ \lambda_2$ 是埃尔米特矩阵 $\boldsymbol{A} \in C^{n \times n}$ 的相异特征值，即

$$\boldsymbol{A}\boldsymbol{x}_1=\lambda_1 \boldsymbol{x}_1,\ \boldsymbol{A}\boldsymbol{x}_2=\lambda_2 \boldsymbol{x}_2,\ \boldsymbol{x}_1,\ \boldsymbol{x}_2 \neq \boldsymbol{\theta} \tag{50}$$

根据定理 7，因 $\lambda_1,\ \lambda_2$ 是实数，由第 2 式得

$$\boldsymbol{x}_2^* \boldsymbol{A}=\lambda_2 \boldsymbol{x}_2^* \tag{51}$$

将(50)式中第 1 式左乘以 $\boldsymbol{x}_2^*$，(51)式右乘以 $\boldsymbol{x}_1$，相减后得

$$\lambda_1 \boldsymbol{x}_2^* \boldsymbol{x}_1=\lambda_2 \boldsymbol{x}_2^* \boldsymbol{x}_1$$

$$(\lambda_1-\lambda_2) \boldsymbol{x}_2^* \boldsymbol{x}_1=0 \tag{52}$$

因 $\lambda_1 \neq \lambda_2$，则 $\boldsymbol{x}_2^* \boldsymbol{x}_1=0$，即 $\boldsymbol{x}_2$ 和 $\boldsymbol{x}_1$ 正交.

证明完毕

［定理 9］ 埃尔米特矩阵是单纯矩阵.

* 原文误为(31)式．——译者注

(证明) 若 $\boldsymbol{A}$ 不是单纯矩阵,如讨论若当标准形时所述,至少对于一个特征值,有使

$$(\lambda \boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}\neq \boldsymbol{0} \tag{53}$$

$$(\lambda \boldsymbol{I}-\boldsymbol{A})^2\boldsymbol{x}=\boldsymbol{0} \tag{54}$$

成立的 $\boldsymbol{x}$ 存在. 设 $\boldsymbol{A}$ 为埃尔米特矩阵, 由定理 7, 因 λ 是实数, 则 $\boldsymbol{B}\triangleq\lambda\boldsymbol{I}-\boldsymbol{A}$ 也是埃尔米特矩阵. 考虑一下使(54)式成立,换言之,满足

$$\boldsymbol{B}^2\boldsymbol{x}=\boldsymbol{0}. \tag{55}$$

的任意向量 $\boldsymbol{x}$. 上式左乘以 $\boldsymbol{x}^*$, 利用 $\boldsymbol{B}=\boldsymbol{B}^*$, 得

$$\boldsymbol{x}^*\boldsymbol{B}^2\boldsymbol{x}=\boldsymbol{x}^*\boldsymbol{B}^*\boldsymbol{B}\boldsymbol{x}=(\boldsymbol{B}\boldsymbol{x})^*(\boldsymbol{B}\boldsymbol{x})=0 \tag{56}$$

由内积公理

$$\boldsymbol{B}\boldsymbol{x}=(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0} \tag{57}$$

因而不存在同时满足(53), (54)式的向量 $\boldsymbol{x}$.

证明完毕

当两个矩阵 $\boldsymbol{A}$, $\boldsymbol{B}$ 对于某个正则矩阵 $\boldsymbol{T}$ 满足

$$\boldsymbol{A}=\boldsymbol{T}^{-1}\boldsymbol{B}\boldsymbol{T} \tag{58}$$

的关系时,称为 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 相似. 其中,当 $\boldsymbol{T}$ 是酉阵时, 特别称为 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 酉相似. 当 $\boldsymbol{T}$ 是正交矩阵时,称为正交相似. 若为酉(正交)相似,则 $\boldsymbol{A}=\boldsymbol{T}^*\boldsymbol{B}\boldsymbol{T}(\boldsymbol{A}=\boldsymbol{T}^T\boldsymbol{B}\boldsymbol{T})$ 成立.

根据定理 8, 与埃尔米特矩阵相异特征值对应的特征向量正交. 因此,若所有特征值均相异,则显然 n 个特征向量中的任意两个都正交. 既使在有相同特征值存在的情况下,使用后面将要讲过的克兰姆-施密特正交化方法, 也可以使任意两个特征向量正交. 而且,若 $\boldsymbol{x}$ 是特征向量, 因为 $\alpha\boldsymbol{x}$, $\alpha\in F$ 也是特征向量, 则适当地规格化后可以选择 n 个满足条件

$$\boldsymbol{x}_i^*\boldsymbol{x}_j=\delta_{ij}\quad i,\ j=1,\ \cdots,\ n \tag{59}$$

的特征向量 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$. 这时,称为 $\boldsymbol{x}_i(i=1,\cdots,n)$ 组成规格化正交集(orthonomal set).

利用系 17 和(10)式可以写成

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{x}_1^*\\ \vdots\\ \boldsymbol{x}_n^*\end{bmatrix}\begin{bmatrix}\lambda_1 & & \boldsymbol{0}\\ & \ddots & \\ \boldsymbol{0} & & \lambda_n\end{bmatrix}[\boldsymbol{x}_1\ \cdots\ \boldsymbol{x}_n]. \tag{60}$$

设 $$\boldsymbol{T}\triangleq[\boldsymbol{x}_1\ \cdots\ \boldsymbol{x}_n],\ \boldsymbol{\Lambda}=\begin{bmatrix}\lambda_1 & & \boldsymbol{0}\\ & \ddots & \\ \boldsymbol{0} & & \lambda_n\end{bmatrix},$$

因 $\boldsymbol{x}_i$ 是规格化正交集的,则 $\boldsymbol{T}^*=\boldsymbol{T}^{-1}$, 即 $\boldsymbol{T}$ 是酉阵. (60)式变成下列关系

$$\boldsymbol{A}=\boldsymbol{T}\boldsymbol{\Lambda}\boldsymbol{T}^*.$$

即有下面定理.

[定理 10] 埃尔米特矩阵与以其特征值为主对角线上元素的对角矩阵酉相似. 反之,若与实对角矩阵酉相似,则是埃尔米特矩阵.

[系 18] 实对称矩阵与以其特征值为主对角线上元素的对角矩阵正交相似. 反之,实矩阵若与实对角矩阵正交相似,则是对称矩阵.

利用该结果,可以推导出关于实对称矩阵的几个性质.

[系 19] 实对称矩阵,当且仅当其所有特征值均为正(非负)时,是正定(准正定)的.

（证明） 对于二次型 $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$, $\boldsymbol{A}\in R^{n\times n}$; $\boldsymbol{x}\in R^n$ 有

$$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}=\boldsymbol{x}^T\boldsymbol{T}^T\boldsymbol{\Lambda}\boldsymbol{T}\boldsymbol{x}=\boldsymbol{y}^T\boldsymbol{\Lambda}\boldsymbol{y} \quad (\boldsymbol{y}\triangleq\boldsymbol{T}\boldsymbol{x}).$$

因 $\det\boldsymbol{T}\neq 0$, 则

$$\boldsymbol{x}=\boldsymbol{0}\Leftrightarrow\boldsymbol{y}=\boldsymbol{0}$$

因此

$$\boldsymbol{A}\text{ 正定}\Leftrightarrow\text{对于 }{}^\forall\boldsymbol{x}\neq\boldsymbol{0},\ \boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}>0\Leftrightarrow\text{对于 }{}^\forall\boldsymbol{y}\neq\boldsymbol{0},$$
$$\boldsymbol{y}^T\boldsymbol{\Lambda}\boldsymbol{y}>0\Leftrightarrow\boldsymbol{\Lambda}\text{ 正定}\Leftrightarrow\text{所有特征值均为正}.$$

若令 $\boldsymbol{y}$ 等于 $\boldsymbol{e}_i(i=1,\ \cdots,\ n)$, 则立即可确认最后的等价性.

对于准正定也是一样.

证明完毕

［系 20］ 设 $\boldsymbol{B}$ 为任意实矩阵,则 $\boldsymbol{A}\triangleq\boldsymbol{B}^T\boldsymbol{B}$ 是准正定的. 反之,若 $\boldsymbol{A}$ 为准正定的,则利用某个实矩阵 $\boldsymbol{B}$ 可以表示成 $\boldsymbol{A}=\boldsymbol{B}^T\boldsymbol{B}$, $\operatorname{rank}\boldsymbol{B}$ 等于 $\boldsymbol{A}$ 的正特征值个数.

（问题 5） 试证明［系 20］.

如［系19］中所述,实对称矩阵或复数埃尔米特矩阵正定(准正定)的充分必要条件是,所有特征值均为正(非负)的.

大家知道,直接根据矩阵的元素判别其是否正定的有下面的西勒维斯特准则(Sylvester's criterion). 这里不予证明,只介绍其结果[25].

以前,我们用在矩阵 $A\in C^{n\times n}$ 中消去若干行和列的方式定义过子式(参看 P-I-2 中第 24 页). 那时在消去第 j 行的同时也消去第 j 列;留下第 i 行的同时也留下第 i 列,按照这种规则作成的子式,即

$$\det\boldsymbol{A}\begin{pmatrix} i_1, & i_2, & \cdots, & i_r \\ i_1, & i_2, & \cdots, & i_r \end{pmatrix} \quad 1\leqslant i_1<i_2\cdots<i_r\leqslant n \quad 1\leqslant r\leqslant n$$

称为 $\boldsymbol{A}$ 的主子式(principal minor).

（西勒维斯特准则）

埃尔米特矩阵 $\boldsymbol{A}=[a_{ij}]\in C^{n\times n}$

（i） $\boldsymbol{A}$ 正定的充分必要条件是,原矩阵 $\boldsymbol{A}$ 左上角的 n 个主子式是正的,即

$$\det\boldsymbol{A}\begin{pmatrix} 1, & 2, & \cdots, & r \\ 1, & 2, & \cdots, & r \end{pmatrix}>0 \quad 1\leqslant r\leqslant n \tag{61}$$

（ii） $\boldsymbol{A}$ 准正定的充分必要条件是,所有主子式皆非负,即

$$\det\boldsymbol{A}\begin{pmatrix} i_1, & i_2, & \cdots, & i_r \\ i_1, & i_2, & \cdots, & i_r \end{pmatrix}\geqslant 0 \quad 1\leqslant i_1<i_2\cdots<i_r\leqslant n \quad 1\leqslant r\leqslant n \tag{62}$$

（注意）

在这里,准正定准则往往容易用错. 这个准则不只是将(61)式的 >0 换成 $\geqslant 0$. 例如,对于

$$\boldsymbol{A}=\begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}$$

$a_{11}\geqslant 0$, $\det\boldsymbol{A}\geqslant 0$; 但显然 $\boldsymbol{A}$ 不是准正定的.

正规阵

设 $A\in C^{n\times n}$ 为埃尔米特矩阵,因 $\boldsymbol{A}^*=\boldsymbol{A}$, 当然

$$A^*A=AA^*$$

成立. 但是反过来, 满足上式的 A 不一定都是埃尔米特矩阵. 一般, 当 $A\in C^{n\times n}$ 满足 $A^*A=AA^*$ 时, A 称为正规阵(normal matrix)(参看 B-I-5 中问题 5). 而且, 当满足

$$A^TA=AA^T$$

时, A 称为实正规阵(real normal matrix). 以下性质很容易证明[23].

(i) 埃尔米特矩阵、斜埃尔米特矩阵及酉阵都是正规阵.

(ii) 实对称、实斜对称及实正交阵都是实正规阵.

(iii) 正规阵, 当且仅当其所有特征值皆为实数时, 是埃尔米特矩阵.

(iv) 正规阵, 当且仅当其所有特征值的实部皆为 0 时, 是斜埃尔米特矩阵.

(v) 正规阵, 当且仅当其所有特征值的绝对值皆为 1 时, 是酉阵.

(vi) 在实正规阵和对称、斜对称、实正交矩阵之间, 也有和(iii)—(v)相同的关系.

其它关于埃尔米特矩阵, 有下列将定理 8, 9, 10 等一般化的性质. 这些性质的证明比较复杂, 可以参阅参考书[6,10,25], 这里省略.

(vii) 与正规阵相异特征值对应的特征向量正交.

(viii) 正规阵是单纯矩阵.

(ix) $A\in C^{n\times n}$ 是正规阵 $\Leftrightarrow A$ 与对角矩阵酉相似.

(x) 设 A 为正规阵, 则 $Ax=\lambda x\Leftrightarrow A^*x=\bar{\lambda}x$

对于实正规阵, 也有同样性质成立.

两个二次型的同时对角变换

对于两个实对称矩阵 $A\in R^{n\times n}$, $B\in R^{n\times n}$, 考虑其二次型

$$Q_1=x^TAx,\ Q_2=x^TBx\quad x\in R^n.$$

我们来证明, 若 A 或 B 有一个是正定的, 例如 A 正定, 则有适当的正则矩阵 $T\in R^{n\times n}$ 存在, 按变量变换 $x=Ty$, 对新变量可以表示成仅包含二乘项的二次型

$$Q_1=y^Ty,\ Q_2=y^T\Lambda y,\ \Lambda=\text{对角矩阵}.$$

若 A 正定, 由[系 20], 则有满足

$$A=P^TP,\ P\in R^{n\times n}$$

的正则矩阵 P 存在. 利用 P 定义

$$\widetilde{B}\triangleq(P^{-1})^TBP^{-1}\in R^{n\times n}.$$

因 B 是实对称的, 则 $\widetilde{B}$ 也是实对称矩阵. 利用[系 18], 将 $\widetilde{B}$ 相似变换成实对角矩阵 Λ, 即由正交矩阵 Q 表示成

$$\widetilde{B}=Q^{-1}\Lambda Q,\quad Q^{-1}=Q^T.$$

现令 $T\triangleq(QP)^{-1}$, 则

$$T^TBT=(Q^{-1})^T(P^{-1})^TBP^{-1}Q^{-1}=(Q^{-1})^T\widetilde{B}Q^{-1}=\Lambda$$

$$T^TAT=(Q^{-1})^T(P^{-1})^TAP^{-1}Q^{-1}=(Q^{-1})^TQ^{-1}=I_n.$$

由此可以证明, 若令 $x=Ty$, 则

$$Q_1=x^TAx=y^TT^TATy=y^Ty$$

$$Q_2=x^TBx=y^TT^TBTy=y^T\Lambda y.$$

(注意)

(i) 一般，$\boldsymbol{P}$ 不是正交矩阵. 因此，$\boldsymbol{T}\triangleq(\boldsymbol{QP})^{-1}$ 也不是正交矩阵、$\boldsymbol{\Lambda}$ 主对角线上的元素(是 $\widetilde{\boldsymbol{B}}$ 的特征值)不是 $\boldsymbol{B}$ 的特征值.

$\widetilde{\boldsymbol{B}}$ 的特征值是 $\det[\lambda\boldsymbol{I}-\widetilde{\boldsymbol{B}}]=0$ 的根，它也是 $\det[\lambda\boldsymbol{A}-\boldsymbol{B}]=0$ 的根或 $\boldsymbol{BA}^{-1}$ 的特征值. 由此可见，例如在 $Q_1=1$ 的约束下，Q_2 的最大值(最小值)是 $\lambda_{\max}(\boldsymbol{BA}^{-1})(\lambda_{\min}(\boldsymbol{BA}^{-1}))$.

(ii) 在 $\boldsymbol{A}$, $\boldsymbol{B}$ 是埃尔米特矩阵且其中有一个是正定阵的情况下，也可以得到同样的结论. 关于两个二次型同时对角化的结论，归纳后便是下列定理.

[定理 11] 设 $\boldsymbol{A}, \boldsymbol{B}\in C^{n\times n}$ 是埃尔米特矩阵，$\boldsymbol{A}$ 是正定矩阵. 这时，有同时满足下列关系

$$\boldsymbol{T}^*\boldsymbol{BT}=\boldsymbol{\Lambda}=\text{实对角矩阵},\ \boldsymbol{T}^*\boldsymbol{AT}=\boldsymbol{I}_n$$

的正则矩阵 $\boldsymbol{T}\in C^{n\times n}$ 存在.

[系 21] 设 $\boldsymbol{A}, \boldsymbol{B}\in R^{n\times n}$ 是对称矩阵，$\boldsymbol{A}$ 是正定矩阵. 这时，有同时满足下列关系

$$\boldsymbol{T}^T\boldsymbol{BT}=\boldsymbol{\Lambda}=\text{实对角矩阵},\ \boldsymbol{T}^T\boldsymbol{AT}=\boldsymbol{I}_n$$

的正则矩阵 $\boldsymbol{T}\in R^{n\times n}$ 存在.

由[系 21]可得出下面结果.

[系 22] 设实对称矩阵 $\boldsymbol{A}, \boldsymbol{B}\in R^{n\times n}$ 均为正定矩阵. 当 $\boldsymbol{x}^T\boldsymbol{Ax}\geqslant\boldsymbol{x}^T\boldsymbol{Bx}$, $^\forall\boldsymbol{x}\in R^n$ 成立(即 $\boldsymbol{A}-\boldsymbol{B}$ 是准正定)时，我们写成 $\boldsymbol{A}\geqslant\boldsymbol{B}$. 这时，$\boldsymbol{0}<\boldsymbol{A}^{-1}\leqslant\boldsymbol{B}^{-1}$.

(证明) 因 $\boldsymbol{B}$ 是正定矩阵，由[系 21]，则有同时满足 $\boldsymbol{T}^T\boldsymbol{BT}=\boldsymbol{I}_n$, $\boldsymbol{T}^T\boldsymbol{AT}=\boldsymbol{\Lambda}=$对角矩阵的正则矩阵 $\boldsymbol{T}\in R^{n\times n}$ 存在. 由于 $\boldsymbol{A}\geqslant\boldsymbol{B}$，则 $\boldsymbol{\Lambda}\geqslant\boldsymbol{I}_n$. 即 $\boldsymbol{\Lambda}-\boldsymbol{I}_n\geqslant\boldsymbol{0}$. 因 $\boldsymbol{\Lambda}$ 和 $\boldsymbol{I}_n$ 是对角矩阵，故这意味着 $\boldsymbol{\Lambda}^{-1}\leqslant\boldsymbol{I}_n$. 因

$$\boldsymbol{A}^{-1}=\boldsymbol{T}^{-1}\boldsymbol{\Lambda}^{-1}(\boldsymbol{T}^{-1})^T,\ \boldsymbol{I}_n=\boldsymbol{T}^{-1}\boldsymbol{B}^{-1}(\boldsymbol{T}^{-1})^T$$

则得 $\boldsymbol{A}^{-1}\leqslant\boldsymbol{B}^{-1}$. 显然，$\boldsymbol{0}<\boldsymbol{A}^{-1}$.

特征值的上限和下限

一般，$\boldsymbol{A}\in C^{n\times n}$ 的特征值有 n 个. 因为它们是特征方程式 $\det[\lambda\boldsymbol{I}-\boldsymbol{A}]=0$ 的根，所以它们分布在复平面上 $\boldsymbol{A}$ 的元素确定的有限范围内. 为了确定其正确位置，从原理上讲，当然要求出 n 次多项式的零点，因而是很困难的. 但是要看所讨论的问题，往往既使只知道特征值存在的范围也是很有益的.

例如，$|\lambda(\boldsymbol{A})|$ 的上限，以及在 $\boldsymbol{A}$ 是埃尔米特矩阵情况下的 $\lambda_{\max}(\boldsymbol{A})$, $\lambda_{\min}(\boldsymbol{A})$ 等等，可以不解特征方程式，直接由 $\boldsymbol{A}$ 的元素求得，所以很方便. 以前对于范数及测度，也推导出给出特征值存在范围的几个关系式. 下面将它们归纳起来，并且给出几个新的结果. 除此之外，关于特征值的存在范围还有许多研究成果，请参看文献[22].

[定理 12]

$$\max|\lambda_i(\boldsymbol{A})|\leqslant\|\boldsymbol{A}\| \tag{63}$$

式中 $\|\boldsymbol{A}\|$ 是任意矩阵范数(参看 B-I-5 中定义 2).

(证明) 一般，若令 $\boldsymbol{A}$ 的若当标准形为 $\boldsymbol{J}$，则 $\boldsymbol{A}^k\to\boldsymbol{0}$, $k\to\infty\Leftrightarrow\boldsymbol{J}^k\to\boldsymbol{0}$, $k\to\infty\Leftrightarrow|\lambda_i(\boldsymbol{A})|<1$, $i=1, 2, \cdots, n$. 设 $\boldsymbol{B}\triangleq\boldsymbol{A}(\varepsilon+\|\boldsymbol{A}\|)^{-1}$, $\varepsilon=$任意正实数，因 $\|\boldsymbol{B}\|=\|\boldsymbol{A}\|(\varepsilon+\|\boldsymbol{A}\|)^{-1}<1$，则 $\boldsymbol{B}^k\to\boldsymbol{0}$, $k\to\infty$.

由此，$|\lambda_i(\boldsymbol{B})|=|\lambda_i(\boldsymbol{A})|(\varepsilon+\|\boldsymbol{A}\|)^{-1}<1$，即 $|\lambda_i(\boldsymbol{A})|<\varepsilon+\|\boldsymbol{A}\|$. 因 $\varepsilon>0$ 是任意的，则 $|\lambda_i(\boldsymbol{A})|\leqslant\|\boldsymbol{A}\|$.

(注意)

(i) 若为埃尔米特矩阵，当使用$\|A\|_2$时，则(63)式中的等号成立. 亦即，由B-I-5中例5，$\|A\|_2=\max_i[\lambda_i(A^*A)]^{1/2}$. 而且由[系7]，$\lambda_i(A^*A)=\lambda_i(A^2)=[\lambda_i(A)]^2$，因此$\|A\|_2=\max_i|\lambda_i(A)|$.

(ii) 对于和向量范数相容的矩阵范数，可以由(47)式得到更简单的关系，即$Ax=\lambda x \Rightarrow |\lambda|\cdot\|x\|=\|Ax\|\leqslant\|A\|\cdot\|x\|\Rightarrow|\lambda|\leqslant\|A\|$.

[定理13] 对于由向量范数导出的矩阵范数$\|A\|$及由其确定的测度$\mu(A)$，下列关系成立.

(i) $-\|A\|\leqslant-\mu(-A)\leqslant\mathrm{Re}\lambda_i(A)\leqslant\mu(A)\leqslant\|A\|,\ i=1,\ 2,\ \cdots,\ n$ (64)

(ii) $\min_i|\lambda_i(A)|\geqslant\max\{-\mu(-A),\ -\mu(A)\}$. (65)

根据B-I-5中矩阵测度的性质(iv)，(vi)，(vii)，该定理很容易证明.

(注意)

(i) 若A是埃尔米特矩阵，则其特征值是实数. 如利用欧几里得范数，因

$$\|A\|_2=\max_i|\lambda_i(A)|,\ \mu_2(A)=\max_i\lambda_i(A),\ -\mu_2(-A)=\min_i\lambda_i(A),$$

则(48)式意味着，当A是埃尔米特矩阵时，

$$-\max|\lambda_i(A)|\leqslant\min\lambda_i(A)\leqslant\lambda_i(A)\leqslant\max_i\lambda_i(A)\leqslant\max_i|\lambda_i(A)|.$$

(ii) (65)式说明，A或$-A$是行和优势的(row-sum dominant)或列和优势的(column-sum dominant)，是A为正则矩阵的充分条件.

(iii) 若令$B\triangleq-jA\ (j^2=-1)$，因$\mathrm{Re}\,\lambda_i(B)=\mathrm{Im}\,\lambda_i(A)$，则由(64)式可以得到对于$A$的特征值虚部的不等式

$$-\mu(jA)\leqslant\mathrm{Im}\,\lambda_i(A)\leqslant\mu(-jA).$$

使用测度$\mu_2(\cdot)$时，由(64)式和上面结果可见，任意$A\in C^{n\times n}$的特征值，其实部和虚部满足下列不等式.

[系23]

$$\min_i\lambda_i\left(\frac{A+A^*}{2}\right)\leqslant\mathrm{Re}\,\lambda_i(A)\leqslant\max_i\lambda_i\left(\frac{A+A^*}{2}\right)$$

$$\min_i\lambda_i\left(\frac{A-A^*}{2j}\right)\leqslant\mathrm{Im}\,\lambda_i(A)\leqslant\max_i\lambda_i\left(\frac{A-A^*}{2j}\right)$$

[定理14] (Geršgorin定理)

在复平面z上考虑$2n$个圆

$$R_i\triangleq\left\{z;\ |z-a_{ii}|\leqslant\sum_{\substack{j=1\\j\neq i}}^{n}|a_{ij}|\right\}$$

$$\left(=以\ a_{ii}\ 为圆心，半径\ r_i\triangleq\sum_{\substack{j=1\\j\neq i}}^{n}|a_{ij}|\ 的圆\right)\quad i=1,\ 2,\ \cdots,\ n$$

及

$$S_j\triangleq\left\{z;\ |z-a_{jj}|\leqslant\sum_{\substack{i=1\\i\neq j}}^{n}|a_{ij}|\right\}$$

$$\left(=以\ a_{jj}\ 为圆心，半径\ S_j\triangleq\sum_{\substack{i=1\\i\neq j}}^{n}|a_{ij}|\ 的圆\right)\quad j=1,\ 2,\ \cdots,\ n$$

$\boldsymbol{A}$ 的全部特征值都存在于 z 平面上的区域 T

$$T=\left(\bigcup_{i=1}^{n}R_i\right)\cap\left(\bigcup_{j=1}^{n}S_i\right)$$

(R_i 的和集与 S_i 的和集的共同部分)

(证明) 设与 $\lambda(\boldsymbol{A})$ 对应的特征向量为 $\boldsymbol{x}=(x_1, x_2, \cdots, x_n)^T$. 因 $\boldsymbol{x}\neq\boldsymbol{\theta}$, 则至少有一个分量不是 0. 设 $\max\limits_{k}|x_k|=|x_i|\neq 0$, 则

$$\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}\Leftrightarrow\sum_{j=1}^{n}a_{ij}x_j=\lambda x_i,\ i=1,\ 2,\ \cdots,\ n$$

由此

$$|\lambda-a_{ii}|\leqslant\sum_{\substack{j=1\\j\neq i}}^{n}|a_{ij}|\left|\frac{x_j}{x_i}\right|\leqslant\sum_{\substack{j=1\\j\neq i}}^{n}|a_{ij}| \tag{66}$$

因为对于任意一个 $i(1\leqslant i<n)$ (66)式都必须成立, 所以 $\lambda(\boldsymbol{A})$ 必须属于集合 R_i 的和集. 而且, 因 $\boldsymbol{A}^T$ 的特征值与 $\boldsymbol{A}$ 的特征值一致, 对于 $\boldsymbol{A}^T$ 使用上面结果, 则 $\lambda(\boldsymbol{A})$ 也必须属于 S_i 的和集. 因此, 必须同时属于 $\bigcup\limits_{i=1}^{n}R_i$ 和 $\bigcup\limits_{j=1}^{n}S_j$. 证明完毕

[定理 15] 设埃尔米特矩阵 $\boldsymbol{A}\in C^{n\times n}$ 的特征值为 $\lambda_1\leqslant\lambda_2\leqslant\cdots\leqslant\lambda_n$, 对应的规格化正交集的特征向量为 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$. 考虑对 $\boldsymbol{y}\neq\boldsymbol{\theta}$ 定义的 $\boldsymbol{y}\in C^n$ 的函数[称为瑞利商 (Rayleigh guotient)]

$$F(\boldsymbol{y})=\frac{\boldsymbol{y}^*\boldsymbol{A}\boldsymbol{y}}{\boldsymbol{y}^*\boldsymbol{y}} \tag{67}$$

这时

$$\lambda_1=F(\boldsymbol{x}_1)\leqslant F(\boldsymbol{y})\leqslant F(\boldsymbol{x}_n)=\lambda_n,\ \forall\boldsymbol{y}\in C^n,\ \boldsymbol{y}\neq\boldsymbol{\theta} \tag{68}$$

(证明) 设以特征向量 $\boldsymbol{x}_i$ 为第 i 列的矩阵

$$\boldsymbol{T}\triangleq[\boldsymbol{x}_1,\ \boldsymbol{x}_2,\ \cdots,\ \boldsymbol{x}_n]\quad \boldsymbol{y}\triangleq\boldsymbol{T}\boldsymbol{z}$$

由定理 10, 因 $\boldsymbol{T}^*\boldsymbol{A}\boldsymbol{T}=\boldsymbol{\Lambda}=\mathrm{diag}(\lambda_1,\ \lambda_2,\ \cdots,\ \lambda_n)$, 则

$$\min_{y}\frac{\boldsymbol{y}^*\boldsymbol{A}\boldsymbol{y}}{\boldsymbol{y}^*\boldsymbol{y}}=\min_{z}\frac{\boldsymbol{z}^*\boldsymbol{\Lambda}\boldsymbol{z}}{\boldsymbol{z}^*\boldsymbol{z}}=\frac{\boldsymbol{e}_1^T\boldsymbol{\Lambda}\boldsymbol{e}_1}{\boldsymbol{e}_1^T\boldsymbol{e}_1}=\lambda_1$$

$$\max_{y}\frac{\boldsymbol{y}^*\boldsymbol{A}\boldsymbol{y}}{\boldsymbol{y}^*\boldsymbol{y}}=\max_{z}\frac{\boldsymbol{z}^*\boldsymbol{\Lambda}\boldsymbol{z}}{\boldsymbol{z}^*\boldsymbol{z}}=\frac{\boldsymbol{e}_n^T\boldsymbol{\Lambda}\boldsymbol{e}_n}{\boldsymbol{e}_n^T\boldsymbol{e}_n}=\lambda_n$$

式中 $\boldsymbol{e}_1\triangleq(1,\ 0,\ \cdots,\ 0)^T$, $\boldsymbol{e}_n\triangleq(0,\ 0,\ \cdots,\ 1)^T$. 对于 $\boldsymbol{z}=\boldsymbol{e}_1$, $\boldsymbol{z}=\boldsymbol{e}_n$, 因 $\boldsymbol{y}=\boldsymbol{x}_1$, $\boldsymbol{y}=\boldsymbol{x}_n$, 则

$$F(\boldsymbol{x}_1)=\lambda_1,\ F(\boldsymbol{x}_n)=\lambda_n$$

证明完毕

(问题 6) 试证明酉阵的特征值分布在复平面的单位圆上.

(问题 7) 试证明斜对称矩阵的特征值分布在虚轴上.

(问题 8) 设 $\boldsymbol{A}$, $\boldsymbol{B}$ 均为埃尔米特矩阵, 而且 $\boldsymbol{B}$ 又是正定矩阵. 试证明, 这时 $\lambda(\boldsymbol{A}\boldsymbol{B}^{-1})$ 是实数,

$$\max_{x}\frac{\boldsymbol{x}^*\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}^*\boldsymbol{B}\boldsymbol{x}}=\lambda_{\max}(\boldsymbol{A}\boldsymbol{B}^{-1}),$$

$$\min_{x}\frac{\boldsymbol{x}^*\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}^*\boldsymbol{B}\boldsymbol{x}}=\lambda_{\min}(\boldsymbol{A}\boldsymbol{B}^{-1})$$

成立.

由矩阵 A, B 生成的各种矩阵的特征值及有关的矩阵不等式

这里将要叙述,由矩阵 A, B 生成的各种矩阵的特征值,以及与其有关的矩阵不等式的若干结果. 例如,讨论 $\lambda(A)$, $\lambda(B)$ 与 $\lambda(A+B)$ 的关系,若 $A>B$, 则 $A^2>B^2$ 也成立等问题. 更详细的内容可参阅文献[22, 62, 63]. 必须注意,在考虑特征值及矩阵不等式时,直观地推断往往会得出错误的结果[64].

首先是以前已讲过的两个结果:

(i) 直和($A\in C^{n\times n}$, $B\in C^{m\times m}$)

$$A\dot{+}B\triangleq\begin{bmatrix}A & 0\\ 0 & B\end{bmatrix}\in C^{(n+m)\times(n+m)}$$

的特征值包含 A 及 B 的特征值的全体,除此之外都不是.

(ii) 克罗内克积和克罗内克和($A\in C^{n\times n}$, $B\in C^{m\times m}$)

$A\otimes B$ 的特征值是 A 及 B 特征值之积, $A\otimes I_m+I_n\otimes B$ 的特征值可以用 A 及 B 特征值之和给出,即

$$\lambda_i(A\otimes B)=\lambda_k(A)\lambda_l(B)$$
$$\lambda_i(A\otimes I_m+I_n\otimes B)=\lambda_k(A)+\lambda_l(B)$$
$$k=1,\ 2\cdots,\ n;\ l=1,\ 2,\ \cdots,\ m$$

其次,我们来看一下由 A, B 生成的最简单的矩阵 $A+B$.

(iii) 和 $A+B$($A\in C^{n\times n}$, $B\in C^{n\times n}$)

首先,由范数和测度的性质(B-I-5),一般关于 $A+B$ 的特征值,

$$\max_i|\lambda_i(A+B)|\leqslant\|A\|+\|B\| \tag{1}$$

及

$$-\mu(-A)-\mu(-B)\leqslant\mathrm{Re}[\lambda_i(A+B)]\leqslant\mu(A)+\mu(B) \tag{2}$$

成立. 在这里,若利用由向量的欧几里得范数导出的矩阵范数 $\|\cdot\|_2$ 和测度 $\mu_2(\cdot)$, 则(1), (2)式又可写成

$$\max_i|\lambda_i(A+B)|\leqslant\lambda_{\max}^{1/2}(A^*A)+\lambda_{\max}^{1/2}(B^*B) \tag{1$'$}$$

$$\lambda_{\min}\left(\frac{A^*+A}{2}\right)+\lambda_{\min}\left(\frac{B^*+B}{2}\right)\leqslant\mathrm{Re}[\lambda_i(A+B)]$$
$$\leqslant\lambda_{\max}\left(\frac{A^*+A}{2}\right)+\lambda_{\max}\left(\frac{B^*+B}{2}\right) \tag{2$'$}$$

若 A, B 均为埃尔米特矩阵,则(1)′, (2)′式可分别表示成

$$\max_i|\lambda_i(A+B)|\leqslant\max_i|\lambda_i(A)|+\max_i|\lambda_i(B)| \tag{1$''$}$$

$$\lambda_{\min}(A)+\lambda_{\min}(B)\leqslant\lambda_i(A+B)\leqslant\lambda_{\max}(A)+\lambda_{\max}(B) \tag{2$''$}$$

一般,若 A, B 不是埃尔米特矩阵,则(1)″, (2)″式不成立. 这意味着,对于(1)″式来说,例如既使 A 和 B 的特征值在单位圆内,但是 $A+B$ 的特征值一般不一定都在以原点为中心半径为 2 的圆内. 另外对于(2)″式,既使 A 和 B 的特征值都分布在左半平面上,但是 $A+B$ 的特征值也不一定都在左半平面. 例如. 对于

$$A=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\qquad B=\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix},$$

有 $|\lambda_i(\boldsymbol{A}+\boldsymbol{B})|=1,\ \lambda_i(\boldsymbol{A})=\lambda_i(\boldsymbol{B})=0$.

另外,对于

$$\boldsymbol{A}=\begin{bmatrix}-\frac{1}{2} & 2\\ 0 & -\frac{1}{2}\end{bmatrix}\quad \boldsymbol{B}=\begin{bmatrix}-\frac{1}{2} & 0\\ 2 & -\frac{1}{2}\end{bmatrix},$$

有 $\lambda_i(\boldsymbol{A})=\lambda_i(\boldsymbol{B})=-\frac{1}{2}$, $\lambda_i(\boldsymbol{A}+\boldsymbol{B})=1,\ -3$. (1)″, (2)″ 两个不等式均不成立.

(注意)

"利用瑞利商也可以得出(2)″式. 即在 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 均为埃尔米特矩阵的情况下,设对应于 $\lambda_i(\boldsymbol{A}+\boldsymbol{B})$ 的特征向量为 $\boldsymbol{x}_i$, 由定理 15 得

$$\begin{aligned}\lambda_i(\boldsymbol{A}+\boldsymbol{B})&=\frac{\boldsymbol{x}_i^*(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{x}_i}{\boldsymbol{x}_i^*\boldsymbol{x}_i}\leqslant\max_{x}\frac{\boldsymbol{x}^*(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{x}}{\boldsymbol{x}^*\boldsymbol{x}}\leqslant\max_{x}\frac{\boldsymbol{x}^*\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}^*\boldsymbol{x}}+\max_{x}\frac{\boldsymbol{x}^*\boldsymbol{B}\boldsymbol{x}}{\boldsymbol{x}^*\boldsymbol{x}}\\&=\lambda_{\max}(\boldsymbol{A})+\lambda_{\max}(\boldsymbol{B})\end{aligned}$$

同样

$$\begin{aligned}\lambda_i(\boldsymbol{A}+\boldsymbol{B})&\geqslant\min_{x}\frac{\boldsymbol{x}^*(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{x}}{\boldsymbol{x}^*\boldsymbol{x}}\geqslant\min_{x}\frac{\boldsymbol{x}^*\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}^*\boldsymbol{x}}+\min_{x}\frac{\boldsymbol{x}^*\boldsymbol{B}\boldsymbol{x}}{\boldsymbol{x}^*\boldsymbol{x}}\\&=\lambda_{\min}(\boldsymbol{A})+\lambda_{\min}(\boldsymbol{B})\text{”}\end{aligned}$$

在 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是埃尔米特矩阵的情况下,容易证明下列不等式也成立.

$$\lambda_{\max}(\boldsymbol{A}+\boldsymbol{B})\geqslant\max\{\lambda_{\max}(\boldsymbol{A})+\lambda_{\min}(\boldsymbol{B}),\ \lambda_{\min}(\boldsymbol{A})+\lambda_{\max}(\boldsymbol{B})\}\tag{3}$$

$$\lambda_{\min}(\boldsymbol{A}+\boldsymbol{B})\leqslant\min\{\lambda_{\max}(\boldsymbol{A})+\lambda_{\min}(\boldsymbol{B}),\ \lambda_{\min}(\boldsymbol{A})+\lambda_{\max}(\boldsymbol{B})\}\tag{4}$$

例如,(3)式证明如下

$$\begin{aligned}\lambda_{\max}(\boldsymbol{A})&=\mu_2(\boldsymbol{A})=\mu_2(\boldsymbol{A}+\boldsymbol{B}-\boldsymbol{B})\leqslant\mu_2(\boldsymbol{A}+\boldsymbol{B})+\mu_2(-\boldsymbol{B})\\&=\lambda_{\max}(\boldsymbol{A}+\boldsymbol{B})-\lambda_{\min}(\boldsymbol{B})\\\lambda_{\max}(\boldsymbol{B})&=\mu_2(\boldsymbol{B})=\mu_2(\boldsymbol{A}+\boldsymbol{B}-\boldsymbol{A})\leqslant\mu_2(\boldsymbol{A}+\boldsymbol{B})+\mu_2(-\boldsymbol{A})\\&=\lambda_{\max}(\boldsymbol{A}+\boldsymbol{B})-\lambda_{\min}(\boldsymbol{A})\end{aligned}$$

(4)式的不等式也可以同样证明. 如上面注意中所述,当然利用瑞利商也可以得出(3),(4)式.

那么,除了 $\boldsymbol{A}$, $\boldsymbol{B}$ 都是埃尔米特矩阵之外,若再假定 $\boldsymbol{A}$, $\boldsymbol{B}$ 之中有一个,例如 $\boldsymbol{B}$ 是正定(准正定)的, 因 $\lambda_i(\boldsymbol{B})>0(\geqslant0)$(证明和系 19 一样),则由(2)″ 和(3)式可见

当 $\boldsymbol{B}>\mathbf{0}$ 时:
$$\lambda_{\max}(\boldsymbol{A}+\boldsymbol{B})>\lambda_{\max}(\boldsymbol{A})$$
$$\lambda_{\min}(\boldsymbol{A}+\boldsymbol{B})>\lambda_{\min}(\boldsymbol{A})$$

当 $\boldsymbol{B}\geqslant\mathbf{0}$ 时:
$$\lambda_{\max}(\boldsymbol{A}+\boldsymbol{B})\geqslant\lambda_{\max}(\boldsymbol{A})$$
$$\lambda_{\min}(\boldsymbol{A}+\boldsymbol{B})\geqslant\lambda_{\min}(\boldsymbol{A})$$

成立. 即, $\boldsymbol{A}$ 加上 $\boldsymbol{B}$ 后, $\boldsymbol{A}$ 的最大及最小特征值都增加. 实际上我们知道,该性质对于中间的所有特征值也很有意义[63]. 即,若加上正定(准正定)矩阵时,则特征值一致增加.

[定理 16]

"设 $\boldsymbol{A}\in C^{n\times n}$, $\boldsymbol{B}\in C^{n\times n}$ 均为埃尔米特矩阵. 将 $\boldsymbol{A}$, $\boldsymbol{A}+\boldsymbol{B}$ 的特征值按照从大到小的顺序进行编号,即 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n$. 当 $\boldsymbol{B}$ 为准正定时

$$\lambda_i(\boldsymbol{A}+\boldsymbol{B})\geqslant\lambda_i(\boldsymbol{A}),\ i=1,\ 2,\ \cdots,\ n\tag{5}$$

成立．若 $\boldsymbol{B}$ 为正定，则不包含等号的不等式成立．”

(iv) 积 $\boldsymbol{AB}(\boldsymbol{A}\in C^{n\times n},\ \boldsymbol{B}\in C^{n\times n})$

对于积 $\boldsymbol{AB}$，我们感兴趣的是，若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的特征值都在单位圆内，那么 $\boldsymbol{AB}$ 所有的特征值是否也都在单位圆内．当 $\boldsymbol{A}$，$\boldsymbol{B}$ 均为埃尔米特矩阵时，这个推测是正确的．这是因为

$$\max_i|\lambda_i(\boldsymbol{AB})|\leqslant\|\boldsymbol{AB}\|_2\leqslant\|\boldsymbol{A}\|_2\|\boldsymbol{B}\|_2=\max_i|\lambda_i(\boldsymbol{A})|\max_i|\lambda_i(\boldsymbol{B})| \tag{6}$$

成立．但是，在一般情况下，这个推测不成立[65]．下面可以作为一个简单的反面例子．

[例 1][66]

$$\boldsymbol{A}=\gamma\begin{bmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}\alpha & 0\\ 0 & \beta\end{bmatrix}$$

$\boldsymbol{A}$ 的特征值存在于单位圆内的充分必要条件是，$|\gamma|<1$．另外对于 $\boldsymbol{B}$，其充分必要条件是

$$|\alpha\beta|<1,\ \alpha\beta-(\alpha+\beta)\cos\theta+1>0$$
$$\alpha\beta+(\alpha+\beta)\cos\theta+1>0$$

现设 $\gamma=0.5$，$\alpha=10$，$\beta=0.01$，$\cos\theta=0.01$，因上述条件均成立，则 $\boldsymbol{A}$，$\boldsymbol{B}$ 的特征值在单位圆内．但是，因

$$\boldsymbol{AB}=\gamma\begin{bmatrix}\alpha & 0\\ 0 & \beta\end{bmatrix}$$

则 $\boldsymbol{AB}$ 的特征值为 $\alpha\gamma=5$，$\beta\gamma=5\times10^{-3}$．

（注意）

“如已讲过，$\boldsymbol{A}^k\to\boldsymbol{0}$，$k\to\infty$ 的充分必要条件是，$\boldsymbol{A}$ 的特征值都在单位圆内．但是由该例可见，既使 $\boldsymbol{A}_k$，$k=1,\ 2,\cdots$ 的所有特征值都在单位圆内，在 $k\to0$ 时，矩阵积 $\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_k$ 也不一定收敛到 $\boldsymbol{0}$．”

一般，因 $\max\limits_i|\lambda_i(\boldsymbol{AB})|\leqslant\|\boldsymbol{AB}\|\leqslant\|\boldsymbol{A}\|\cdot\|\boldsymbol{B}\|$，若使用 $\|\cdot\|_2$，则对于任意 $\boldsymbol{A}\in C^{n\times n}$，$\boldsymbol{B}\in C^{n\times n}$，

$$\max_i|\lambda_i(\boldsymbol{AB})|\leqslant\max_i\lambda_i^{1/2}(\boldsymbol{A}^*\boldsymbol{A})\max_i\lambda_i^{1/2}(\boldsymbol{B}^*\boldsymbol{B}) \tag{7}$$

成立．

关于矩阵积，在系统理论中经常用的有下面的不等式．

[定理 17][67] 对于任意埃尔米特矩阵 $\boldsymbol{A}\in C^{n\times n}$，$\boldsymbol{B}\in C^{n\times n}$，

$$\lambda_{\min}(\boldsymbol{B})\lambda_i(\boldsymbol{A}^2)\leqslant\lambda_i(\boldsymbol{ABA})\leqslant\lambda_{\max}(\boldsymbol{B})\lambda_i(\boldsymbol{A}^2)\quad i=1,\ 2,\ 3,\ \cdots,\ n \tag{8}$$

成立．其中假定特征值按照 $\lambda_{\max}=\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n=\lambda_{\min}$ 的顺序编号．”

（证明）在

$$\lambda_{\max}(\boldsymbol{B})\boldsymbol{A}^2=\boldsymbol{A}\{\lambda_{\max}(\boldsymbol{B})\boldsymbol{I}-\boldsymbol{B}\}\boldsymbol{A}+\boldsymbol{ABA} \tag{9}$$

中，$\lambda_{\max}(\boldsymbol{B})\boldsymbol{I}-\boldsymbol{B}$ 是准正定的．因此，(9)式右边第一项是准正定的．将定理 16 用于(9)式，得

$$\lambda_i(\boldsymbol{ABA})\leqslant\lambda_{\max}(\boldsymbol{B})\lambda_i(\boldsymbol{A}^2).$$

此外，若将定理 16 用于

$$-\lambda_{\min}(\boldsymbol{B})\boldsymbol{A}^2=\boldsymbol{A}\{\boldsymbol{B}-\lambda_{\min}(\boldsymbol{B})\boldsymbol{I}\}\boldsymbol{A}-\boldsymbol{ABA} \tag{10}$$

则得 $\lambda_i(\boldsymbol{ABA})\geqslant\lambda_{\min}(\boldsymbol{B})\lambda_i(\boldsymbol{A}^2)$．

[系 23][68]

对于任意准正定埃尔米特矩阵 $\boldsymbol{A}\in C^{n\times n}$, $\boldsymbol{B}\in C^{n\times n}$,

$$\lambda_{\min}(\boldsymbol{B})\lambda_{\max}(\boldsymbol{A})\leqslant\lambda_{\min}(\boldsymbol{B})\operatorname{tr}\boldsymbol{A}\leqslant\operatorname{tr}\boldsymbol{AB}\leqslant\lambda_{\max}(\boldsymbol{B})\operatorname{tr}\boldsymbol{A} \tag{11}$$

成立.

(证明) 由

$$\boldsymbol{B}\geqslant\boldsymbol{0}\Rightarrow\lambda_{\min}(\boldsymbol{B})\geqslant 0,\quad \boldsymbol{A}\geqslant\boldsymbol{0}\Rightarrow\operatorname{tr}\boldsymbol{A}=\sum_{i=1}^{n}\lambda_i(\boldsymbol{A})\geqslant\lambda_{\max}(\boldsymbol{A})$$

可以证明(11)式中的第一个不等式.

因 $\boldsymbol{A}^*=\boldsymbol{A}$, 则有某个酉阵 $\boldsymbol{P}$, 可以分解成 $\boldsymbol{A}=\boldsymbol{P}^*\boldsymbol{\Lambda P}$(参照定理 10), 其中 $\boldsymbol{\Lambda}$ 是以 $\lambda_i(\boldsymbol{A})\geqslant 0$ 为主对角线上元素的实对角矩阵. 现设对角线上元素中具有 $\lambda_i^{1/2}(\boldsymbol{A})\geqslant 0$ 的实对角矩阵为 $\boldsymbol{\Lambda}^{1/2}$, 定义 $\boldsymbol{A}^{1/2}\triangleq\boldsymbol{P}^*\boldsymbol{\Lambda}^{1/2}\boldsymbol{P}$. 注意$(\boldsymbol{A}^{1/2})^2=\boldsymbol{A}$, 在定理 17 中,若作 $\boldsymbol{A}\to\boldsymbol{A}^{1/2}$, $\boldsymbol{A}^2\to\boldsymbol{A}$ 的置换,则可得

$$\lambda_{\min}(\boldsymbol{B})\lambda_i(\boldsymbol{A})\leqslant\lambda_i(\boldsymbol{A}^{1/2}\boldsymbol{B}\boldsymbol{A}^{1/2})\leqslant\lambda_{\max}(\boldsymbol{B})\lambda_i(\boldsymbol{A}) \tag{12}$$

因此,考虑到

$$\operatorname{tr}(\boldsymbol{A}^{1/2}\boldsymbol{B}\boldsymbol{A}^{1/2})=\operatorname{tr}\boldsymbol{AB}(=\operatorname{tr}\boldsymbol{BA})$$

可以得出

$$\lambda_{\min}(\boldsymbol{B})\operatorname{tr}\boldsymbol{A}\leqslant\sum_{i=1}^{n}\lambda_i(\boldsymbol{A}^{1/2}\boldsymbol{B}\boldsymbol{A}^{1/2})=\operatorname{tr}\boldsymbol{AB}\leqslant\lambda_{\max}(\boldsymbol{B})\operatorname{tr}\boldsymbol{A}.$$

(问题 9) 设 $\boldsymbol{A}\in C^{n\times n}$, $\boldsymbol{B}\in C^{n\times n}$ 均为正定(准正定)埃尔米特矩阵. 试证明 $\boldsymbol{AB}$ 为正定(准正定)埃尔米特矩阵的充分必要条件是, $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 是可换的.

(v) 矩阵不等式 $\boldsymbol{A}\geqslant\boldsymbol{B}$

埃尔米特矩阵 $\boldsymbol{A}\in C^{n\times n}$, $\boldsymbol{B}\in C^{n\times n}$ 满足 $\boldsymbol{A}-\boldsymbol{B}\geqslant\boldsymbol{0}$(为准正定)时, 表示成$\boldsymbol{A}\geqslant\boldsymbol{B}$.

当 $\boldsymbol{A}\geqslant\boldsymbol{B}$ 时,对 $\boldsymbol{A}=(\boldsymbol{A}-\boldsymbol{B})+\boldsymbol{B}$ 应用定理 16, 则

$$\lambda_i(\boldsymbol{A})=\lambda_i(\boldsymbol{A}-\boldsymbol{B}+\boldsymbol{B})\geqslant\lambda_i(\boldsymbol{B})\quad i=1,\ 2,\ \cdots,\ n \tag{13}$$

成立. 但是其逆一般不成立, 即 $\lambda_i(\boldsymbol{A})\geqslant\lambda_i(\boldsymbol{B})$ 不一定意味着 $\boldsymbol{A}\geqslant\boldsymbol{B}$. 由此可见, 既使 $\boldsymbol{A}\geqslant\boldsymbol{B}$, 但是一般 $\boldsymbol{A}^2\geqslant\boldsymbol{B}^2$ 并不正确.

[例 2][64]

$$\boldsymbol{B}=\begin{bmatrix}4.4 & 2.5\\ 2.5 & 1.5\end{bmatrix}>\boldsymbol{0},\qquad \boldsymbol{A}=\begin{bmatrix}5 & 3\\ 3 & 2\end{bmatrix}>\boldsymbol{0}$$

$$\boldsymbol{A}-\boldsymbol{B}=\begin{bmatrix}0.6 & 0.5\\ 0.5 & 0.5\end{bmatrix}>\boldsymbol{0}$$

$$\lambda_1(\boldsymbol{A})=6.85,\ \lambda_2(\boldsymbol{A})=0.15$$

$$\lambda_1(\boldsymbol{B})=5.84,\ \lambda_2(\boldsymbol{B})=0.06$$

虽然

$$\boldsymbol{A}^2-\boldsymbol{B}^2=\begin{bmatrix}8.39 & 6.25\\ 6.25 & 4.50\end{bmatrix}$$

是正定的,但不是准正定的.

[例 3]

$$\boldsymbol{A}=\begin{bmatrix}1.3 & 0\\ 0 & 1\end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix}1 & 0.2\\ 0.2 & 1\end{bmatrix}$$

$$\lambda_1(\boldsymbol{A})=1.3,\quad \lambda_2(\boldsymbol{A})=1$$
$$\lambda_1(\boldsymbol{B})=1.2,\quad \lambda_2(\boldsymbol{B})=0.8$$
$$\boldsymbol{A}-\boldsymbol{B}=\begin{bmatrix}0.3 & -0.2\\ -0.2 & 0\end{bmatrix}$$
$$\lambda_1(\boldsymbol{A})>\lambda_1(\boldsymbol{B}),\ \lambda_2(\boldsymbol{A})>\lambda_2(\boldsymbol{B})$$

虽然 $\boldsymbol{A}-\boldsymbol{B}$ 是正定的，但它不是准正定的.

我们知道[69]，例 2 的逆，即“若 $\boldsymbol{A}^2\geqslant\boldsymbol{B}^2$，则 $\boldsymbol{A}\geqslant\boldsymbol{B}$”在 $\boldsymbol{A}>\boldsymbol{0}$，$\boldsymbol{B}>\boldsymbol{0}$ 时是正确的. 而且，F. T. 曼将上述结论普遍化，得出如下结果.

[定理 18][70]　当 $\boldsymbol{A}\in C^{n\times n}$，$\boldsymbol{B}\in C^{n\times n}$，$\boldsymbol{S}\in C^{n\times n}$ 都是正定埃尔米特矩阵时，若 $\boldsymbol{ASA}>\boldsymbol{BSB}$，则 $\boldsymbol{A}>\boldsymbol{B}$.

(证明)

$$\boldsymbol{ASA}-\boldsymbol{BSB}>\boldsymbol{0}$$

和

$$\boldsymbol{S}-\boldsymbol{A}^{-1}\boldsymbol{BSBA}^{-1}>\boldsymbol{0} \tag{14}$$

是等价的. (14)式意味着，在线性采样控制系统 $\boldsymbol{x}(k+1)=\boldsymbol{BA}^{-1}\boldsymbol{x}(k)$，$k=0, 1, 2, \cdots$ 中，二次型李亚普诺夫函数 $V(\boldsymbol{x})\triangleq\boldsymbol{x}^*\boldsymbol{Sx}$ 的差分 $V(\boldsymbol{x}(k))-V(\boldsymbol{x}(k+1))=\boldsymbol{x}^*(k)[\boldsymbol{S}-\boldsymbol{A}^{-1}\boldsymbol{BSBA}^{-1}]\boldsymbol{x}(k)$，在 $\boldsymbol{x}(k)\neq\boldsymbol{0}$ 时，始终为正的. 根据以后将要讨论的系统稳定性理论，这时采样控制系统 $\boldsymbol{x}(k+1)=\boldsymbol{BA}^{-1}\boldsymbol{x}(k)$ 渐近稳定. 即，对于任意 $\boldsymbol{x}(0)\in C^n$，为 $\boldsymbol{x}(k)\to\boldsymbol{0}$，$k\to\infty$. 因 $\boldsymbol{x}(k)=(\boldsymbol{BA}^{-1})^k\boldsymbol{x}(0)$，则还意味着 $\max\limits_i|\lambda_i(\boldsymbol{BA}^{-1})|<1$.

也就是说，当(14)式成立时，$\boldsymbol{BA}^{-1}$ 的特征值必须在单位圆内. 因

$$\boldsymbol{A}^{-1/2}\boldsymbol{BA}^{-1}\boldsymbol{A}^{1/2}=\boldsymbol{A}^{-1/2}\boldsymbol{BA}^{-1/2} \tag{15}$$

则 $\boldsymbol{BA}^{-1}$ 和埃尔米特矩阵 $\boldsymbol{A}^{-1/2}\boldsymbol{BA}^{-1/2}$ 相似，因而有相同的特征值. 考虑正定埃尔米特形式

$$\boldsymbol{x}^*(\boldsymbol{I}-\boldsymbol{A}^{-1/2}\boldsymbol{BA}^{-1/2})\boldsymbol{x}$$

若换成 $\boldsymbol{y}=\boldsymbol{A}^{-1/2}\boldsymbol{x}$，则得

$$\boldsymbol{y}^*(\boldsymbol{A}-\boldsymbol{B})\boldsymbol{y}>0,\ \forall\boldsymbol{y}\neq\boldsymbol{0} \tag{16}$$

这意味着 $\boldsymbol{A}>\boldsymbol{B}$.

(注意)

定理 18 的逆，即“$\boldsymbol{A}>\boldsymbol{B}>\boldsymbol{0}\Rightarrow\boldsymbol{ASA}>\boldsymbol{BSB}$, $\forall\boldsymbol{S}>\boldsymbol{0}$”，如对于 $\boldsymbol{S}=\boldsymbol{I}$ 反例所示，一般是不正确的. 但是我们知道[70]，对于某个 $\boldsymbol{S}>\boldsymbol{0}$，该逆成立.

[系 24]　若正定的埃尔米特矩阵 $\boldsymbol{A}\in C^{n\times n}$，$\boldsymbol{B}\in C^{n\times n}$ 是可换的，则 $\boldsymbol{A}>\boldsymbol{B}\Rightarrow\boldsymbol{A}^k>\boldsymbol{B}^k$，$k=2, 3, \cdots$.

(证明)

$$\boldsymbol{A}>\boldsymbol{B}\Leftrightarrow\max_i|\lambda_i(\boldsymbol{BA}^{-1})|<1.$$

$$\max_i|\lambda_i(\boldsymbol{BA}^{-1})|<1\Rightarrow\max_i|\lambda_i(\boldsymbol{BA}^{-1})^2|<1,$$

因 $\boldsymbol{A}$, $\boldsymbol{B}$ 可换，$\boldsymbol{A}^{-1}\boldsymbol{B}=\boldsymbol{BA}^{-1}$，则

$$\Rightarrow\max_i|\lambda_i(\boldsymbol{B}^2\boldsymbol{A}^{-2})|<1\Rightarrow\boldsymbol{A}^2>\boldsymbol{B}^2,$$

以下，利用归纳法得 $\boldsymbol{A}^k>\boldsymbol{B}^k$，$k=2, 3, \cdots$.

以上，证明了矩阵不等式 $A>B$ 的几个有关的结果．最后，我们将这些结果整理列于表1.

表1 $A>B$ 的有关性质 $A\in C^{n\times n}$, $B\in C^{n\times n}$ 均为埃尔米特矩阵

条　件　(i)	条　件　(ii)	(i)⇒(ii)	备　　注
$A\geqslant B$	$\lambda_i(A)\geqslant\lambda_i(B)$ $i=1, 2, \cdots, n$	成　　立	若 $A>B$，则不包含等号
$A>B>0$	$ASA>BSB$ $\forall S=S^*>0$	一般不成立	对于某个 $S>0$ 成立
$A>B>0$	$A^k>B^k$ $k=2, 3, \cdots$	一般不成立	若 $AB=BA$ 成立
$\lambda_i(A)\geqslant\lambda_i(B)$ $i=1, 2, \cdots, n$	$A\geqslant B$	一般不成立	
$ASA>BSB$ $S=S^*>0$, $A>0$, $B>0$	$A>B$	成　　立	

(vi) 矩阵不等式及收敛

假定有某个实数序列 $\{a_n\}$，在各元素之间有 $a_n\geqslant a_{n+1}$, $n=1, 2, \cdots$ 的关系．这时，若 $a_n\geqslant k$, $n=1, 2, \cdots$ 有下限，则 $\lim\limits_{n\to\infty}a_n$ 存在．对于矩阵序列也有同样的结论．

[定理 19]　在实对称矩阵序列 $\{A\}$ 中，当 $A_n\geqslant K$, $A_n\geqslant A_{n+1}$, $n=1, 2, \cdots$ 成立时，$\lim\limits_{n\to\infty}A_n$ 存在．其中 K 是某个固定的实对称矩阵．

(证明)　初步证明如下．因

$$x^TA_1x\geqslant x^TA_2x\geqslant\cdots\geqslant x^TKx,\ \forall x\in R^n$$

将 $x=e_i=(0, 0, 1, \cdots, 0)^T$, $i=1, 2, \cdots, n$ 这样的单位向量代入上式，则可以证明 A_n 主对角线上的元素有极限．其次，对于 $x=(1, 1, 0, \cdots, 0)^T$，可以证明 A_n 的 $(1, 2)$ 元素 $(=(2, 1)$ 元素$)$ 收敛．

以下同样，可以证明 A_n 的全部元素都收敛．

P-I-4 多　项　式

设 F 是一个体，$a_i(i=0,1,\cdots,n)$是 F 的元．当令 λ 为变量时，大家知道

$$f(\lambda)\triangleq a_0+a_1\lambda+\cdots+a_n\lambda^n=\sum_{i=0}^{n}a_i\lambda^i \tag{1}$$

称为变量 λ 的多项式(参看 B-0-1 中例 6)．

若 $a_n\neq 0$，则 n 称为多项式 $f(\lambda)$ 的次数，用 $\deg f$ 表示．若 $a_n=1$，则多项式称为首一多项式(monic polynomial)．

大家知道，两个多项式的和，积可按 B-0-1 中(2)，(3)式定义，并且其次数之间有下列关系

$$\deg(f+g)\leqslant\max(\deg f,\ \deg g) \tag{2}$$

$$\deg f\cdot g=\deg f+\deg g \tag{3}$$

对于多项式 f，$g(g\neq 0)$，选取适当的多项式 q，r，可以满足下式

$$f=gq+r,\ \deg r<\deg g \tag{4}$$

式中 q 称为 f 用 g 除得的商，r 称为它的余式．当 $r=0$，即没有余式时，称为 f 能用 g 除尽(整除)．当 f_1，f_2 都能用 g 除尽，而用比 g 的次数高的任何多项式都不能除尽时，g 称为 f_1，f_2 的最大公约式．当 f_1 和 f_2 的最大公约式的次数为 0 时，称为 f_1 和 f_2 互素．

［定理］[71]

$$\left.\begin{aligned}f_1(\lambda)&=a_0^1+a_1^1\lambda+\cdots+a_n^1\lambda^n\\ f_2(\lambda)&=a_0^2+a_1^2\lambda+\cdots+a_m^2\lambda^m\end{aligned}\right\} \tag{5}$$

互素的充分必要条件是，矩阵 $\boldsymbol{R}=(n+m)\times(n+m)$

$$\boldsymbol{R}\triangleq\begin{bmatrix}a_n^1 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & a_m^2\\ a_{n-1}^1 & a_n^1 & \cdots & 0 & 0 & 0 & 0 & \cdots & a_m^2 & a_{m-1}^2\\ a_{n-2}^1 & a_{n-1}^1 & \cdots & 0 & 0 & 0 & 0 & \cdots & a_{m-1}^2 & a_{m-2}^2\\ \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots\\ 0 & 0 & \cdots & a_0^1 & a_1^1 & a_1^2 & a_0^2 & \cdots & 0 & 0\\ 0 & 0 & \cdots & 0 & a_0^1 & a_0^2 & 0 & \cdots & 0 & 0\end{bmatrix} \tag{6}$$

是正则的．

若 f_1 和 f_2 互素，则有使

$$f_1\cdot p_1+f_2\cdot p_2=1 \tag{7}$$

的多项式 p_1，p_2 存在[9]．

使 $f(\lambda)=0$ 的 λ 值，叫做 $f(\lambda)$ 的零点或根．

特别是，若 F 为复数体 C，则 $f(\lambda)$ 具有 n 个根 $\lambda_1,\cdots,\lambda_n\in C$ 时可以表示成

$$f(\lambda)=a_n(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n) \tag{8}$$

该式称为质因子分解．

B-I-12 矩阵多项式

设 A 是一个体 F 上的 $n\times n$ 阶方阵，$f(\lambda)$ 为 F 上的多项式．在 $f(\lambda)=a_0+a_1\lambda+\cdots+a_n\lambda^n$ 中，将 λ^i 换成 A^i，$\lambda^0=1$ 换成单位矩阵 I，可以作成一个矩阵多项式

$$a_0 I+a_1 A+\cdots+a_n A^n \tag{1}$$

上式常写成 $f(A)$，称为矩阵多项式[1)]．若 $a_n\neq 0$，则 n 称为 $f(A)$ 的次数．

两个矩阵多项式的和，积分别定义成[2)]

$$f(A)+g(A)\triangleq(f+g)(A) \tag{2}$$

$$f(A)\cdot g(A)\triangleq(f\cdot g)(A) \tag{3}$$

注意第二式的右边，因多项式乘法运算是可换的，则可得出

$$f(A)\cdot g(A)=(f\cdot g)(A)=(g\cdot f)(A)=g(A)\cdot f(A) \tag{4}$$

即与矩阵乘法运算不可换无关，有下列性质

（i）对于任意矩阵 A，矩阵多项式的乘法运算始终是可换的．

除了该性质之外，现再来讨论一下矩阵多项式其它几个性质．

（ii）$f(TAT^{-1})=Tf(A)T^{-1}$ (5)

（证明） 显然，$I=TIT^{-1}$，$(TAT^{-1})^2=TAT^{-1}TAT^{-1}=TA^2T^{-1}$．而且一般 $(TAT^{-1})^i=TA^iT^{-1}$．将这些式子代入 $f(TAT^{-1})$，利用分配律即可得到 $Tf(A)T^{-1}$．

(iii) 若进行适当编号，则 $\lambda_i[f(A)]=f[\lambda_i(A)]$．而且，与 A 的 $\lambda_i(A)$ 对应的特征向量，是与 $f(A)$ 的 $f[\lambda_i(A)]$ 对应的特征向量．其逆亦真．

（证明） 和 B-I-11 中系 7 同样．

(iv) 设按 B-I-11 中定理 5 得到单纯矩阵 A 的谱分解为

$$A=\sum_{j=1}^{\sigma}\lambda_i E_j \tag{6}$$

则

$$f(A)=\sum_{j=1}^{\sigma}f(\lambda_i)E_j \tag{7}$$

成立．

（证明） 根据 B-I-11 中定理 5 的性质(ii)，(iii)，由 $A^k=\sum_{j=1}^{\sigma}(\lambda_j)^k E_j$ 显然．

---

1) 设 $f_{ij}(\lambda)\,(i=1,\cdots,m;\ j=1,\cdots,n)$ 均为同一个体 F 上的多项式．这时，常把以 f_{ij} 为元素的 $m\times n$ 矩阵 $F(\lambda)=[f_{ij}(\lambda)]$ 叫做多项式矩阵．因为矩阵多项式与多项式矩阵不同，所以一定不要混淆．设 $p\triangleq\max\limits_{ij}\deg f_{ij}$，使用适当的矩阵序列 $F_i\in F^{m\times n}(i=1,\cdots,p)$，可将 $F(\lambda)$ 写成

$$F(\lambda)=F_0+\lambda F_1+\lambda^2F_2+\cdots+\lambda^pF_p \tag{1$'$}$$

与(1)式比较显然不同．

但是，(1)$'$ 式和 $f(\lambda)$ 具有同样的形式，其区别仅仅是，(1)$'$ 式的系数不是数，而是矩阵，也可以说它是以矩阵为系数的多项式．因此，和由 $f(\lambda)$ 定义 $f(A)$ 一样，将 λ^i 换成 A^i，$\lambda^0=1$ 换成 I，象这样作成的 $F(A)$ 又如何呢？容易看到，当且仅当 $m=n$，A 和 $F_i(i=0,1,\cdots,p)$ 可换时，$F(A)$ 才有意义，这时 $F(A)\in F^{n\times n}$．但是，既使在这种情况下，关于 $f(A)$ 成立的性质(i)～(iv)，对于 $F(A)$ 也不一定成立．

2) 例如 $(f\cdot g)(A)$ 表示，先按照 B-0-1 中(3)式求 $f(\lambda)$ 和 $g(\lambda)$ 的积，然后在所得到的多项式中，将 λ^i 换成 A^i，$\lambda^0=1$ 换成 I．

此外，将矩阵 $\boldsymbol{A}\in F^{n\times n}$ 固定，这时满足

$$f(\boldsymbol{A})=\boldsymbol{0} \tag{8}$$

的多项式 f 称为 $\boldsymbol{A}$ 的零化多项式(annihilating polynomial).

下面我们来证明，$\boldsymbol{A}$ 的特征多项式(参看 B-I-11)是 $\boldsymbol{A}$ 的零化多项式. 这就是所谓的凯莱-哈密顿(Cayley-Hamilton)定理.

[定理 1] 设 $\psi(\lambda)\triangleq\det(\lambda\boldsymbol{I}-\boldsymbol{A})$，则

$$\psi(\boldsymbol{A})=\boldsymbol{0} \tag{9}$$

成立.

(证明) $(\lambda\boldsymbol{I}-\boldsymbol{A})$的余子式(参看 P-I-2)* 最多是 λ 的 $(n-1)$ 次多项式(为何?). 因此，根据上页注解 2，利用某个矩阵序列 $\boldsymbol{B}_i\in F^{n\times n}(i=0,1,\cdots,n-1)$，伴随矩阵 $\operatorname{adj}(\lambda\boldsymbol{I}-\boldsymbol{A})$ 可写成

$$\operatorname{adj}(\lambda\boldsymbol{I}-\boldsymbol{A})=\boldsymbol{B}_0+\lambda\boldsymbol{B}_1+\cdots+\lambda^{n-1}\boldsymbol{B}_{n-1} \tag{10}$$

将该式代入 P-I-2 中(9)式，得

$$\begin{aligned}(\lambda\boldsymbol{I}-\boldsymbol{A})&(\boldsymbol{B}_0+\lambda\boldsymbol{B}_1+\cdots+\lambda^{n-1}\boldsymbol{B}_{n-1})\\&=(\boldsymbol{B}_0+\lambda\boldsymbol{B}_1+\cdots+\lambda^{n-1}\boldsymbol{B}_{n-1})(\lambda\boldsymbol{I}-\boldsymbol{A})=\psi(\lambda)\boldsymbol{I}\end{aligned} \tag{11}$$

由其中第一个等式可以确定，$\boldsymbol{A}$ 和 $\boldsymbol{B}_i$ 是可易的. 再根据上述注解，在上式中将 λ^i 换成 $\boldsymbol{A}^i$，$\lambda^0=1$ 换成 $\boldsymbol{I}$ 后是有定义的. 因此

$$\psi(\boldsymbol{A})\boldsymbol{I}=\psi(\boldsymbol{A})=(\boldsymbol{A}-\boldsymbol{A})(\boldsymbol{B}_0+\boldsymbol{A}\boldsymbol{B}_1+\cdots+\boldsymbol{A}^{n-1}\boldsymbol{B}_{n-1})=\boldsymbol{0}. \tag{12}$$

证明完毕

在这里可以看到，n 次多项式 $\psi(\lambda)$ 是 $\boldsymbol{A}$ 的零化多项式. 但是，根据具体的矩阵，有次数更小的零化多项式. 例如，无论 n 多么大，$\boldsymbol{I}_n$ 具有一次零化多项式 $(1-\lambda)$. 在 $\boldsymbol{A}$ 的零化多项式中，次数最低的叫做 $\boldsymbol{A}$ 的最小多项式(minimal polynomial)，写成 $\psi_m(\lambda)$.

最小多项式的求法后面再讲，先稍介绍一下其性质.

[定理 2] $\boldsymbol{A}$ 的任意零化多项式都可以用 $\boldsymbol{A}$ 的最小多项式整除.

(证明) 如 P-I-4 中所述，对于 $\deg f>\deg\psi_m$ 的任意 f，可以表示成

$$f=\psi_m q+r \quad \deg r<\deg\psi_m \tag{13}$$

假定 f 不能用 ψ_m 除尽，即 $r\neq 0$. 因为若 f 是零化多项式，则 $f(\boldsymbol{A})=\boldsymbol{0}$，而且 $\psi_m(\boldsymbol{A})=\boldsymbol{0}$，则

$$r(\boldsymbol{A})=\boldsymbol{0} \tag{14}$$

必定成立. 因 r 的次数比 ψ_m 的次数小，所以与 ψ_m 是最小多项式相矛盾.

证明完毕

[定理 3] 所给矩阵 $\boldsymbol{A}$ 的首一最小多项式是唯一确定的.

(证明) 设 ψ_m^1, ψ_m^2均为 $\boldsymbol{A}$ 的最小多项式，则其次数相等. 而且，根据定理 2，能互相除尽. 因而，对于某个 $a\in F$，

$$\psi_m^1=a\psi_m^2 \tag{15}$$

成立. 若 ψ_m^1, ψ_m^2 均为首一多项式，则必须 $a=1$.

证明完毕

以下，若提到 $\boldsymbol{A}$ 的最小多项式，即指首一的.

下面对于 $F=C$ 的情况，来求一下最小多项式．首先，利用矩阵多项式的性质(ii)，因 $\boldsymbol{T}$ 是正则的，则

$$f(\boldsymbol{A})=\boldsymbol{0} \Leftrightarrow f(\boldsymbol{TAT}^{-1})=\boldsymbol{0} \tag{16}$$

即下面引理成立．

［引理 1］ 相似矩阵的最小多项式相同．因此，为了求 $\boldsymbol{A}$ 的最小多项式 $\psi_m(\lambda)$，可以来研究和 $\boldsymbol{A}$ 相似的若当标准形 $\boldsymbol{J}$.

$\boldsymbol{J}$ 的特征多项式是

$$\psi=(\lambda-\lambda_1)^{m_1}\cdots(\lambda-\lambda_\sigma)^{m_\sigma} \tag{17}$$

式中 λ_j 是 $\boldsymbol{A}$ 的(因而也是 $\boldsymbol{J}$ 的)相异特征值，m_j 是 λ_j 的代数重复度．根据定理1，2，$\psi(\lambda)$ 可以用最小多项式 $\psi_m(\lambda)$ 整除．因此，ψ_m 必须成

$$\bar{\psi}(\lambda) \triangleq (\lambda-\lambda_1)^{d_1}\cdots(\lambda-\lambda_\sigma)^{d_\sigma} \quad d_j \leqslant m_j \quad j=1,\ 2,\ \cdots,\ \sigma \tag{18}$$

的形式．求 $\psi_m(\lambda)$ 归结为，求满足 $\bar{\psi}(\boldsymbol{A})=\boldsymbol{0}$ 的最小的 d_j.

由 B-I-11 中(19)～(21)式容易看到，由(18)式的 $\bar{\psi}$ 作成的 $\bar{\psi}(\boldsymbol{J})$，是以

$$\boldsymbol{K}_{jk} \triangleq (\boldsymbol{J}_{jk}-\lambda_1\boldsymbol{I})^{d_1}\cdots(\boldsymbol{J}_{jk}-\lambda_\sigma\boldsymbol{I})^{d_\sigma} \tag{19}$$

为主对角线上元素的分块对角矩阵．如 B-I-11 中(36)式所示，因为 $l \neq j$ 时 $(\boldsymbol{J}_{jk}-\lambda_l\boldsymbol{I})$ 是正则的，则

$$\begin{aligned}\bar{\psi}(\boldsymbol{J})=\boldsymbol{0} &\Leftrightarrow \boldsymbol{K}_{jk}=\boldsymbol{0},\ K=1,\ \cdots,\ \alpha_j;\ j=1,\ \cdots,\ \sigma\\ &\Leftrightarrow (\boldsymbol{J}_{jk}-\lambda_j\boldsymbol{I})^{d_j}=\boldsymbol{0},\ K=1,\ \cdots,\ \alpha_j;\ j=1,\ \cdots,\ \sigma\\ &\Leftrightarrow d_j \geqslant \max_K n_{jk} \quad i=1,\ \cdots,\ \sigma\end{aligned} \tag{20}$$

成立．最后的 $\Leftrightarrow$ 是 B-I-11 中(42)式的结果．于是，得到下列定理．

［定理 4］ 矩阵 $\boldsymbol{A}$ 的最小多项式可以用

$$\psi_m(\lambda)=(\lambda-\lambda_1)^{d_1}\cdots(\lambda-\lambda_\sigma)^{d_\sigma} \tag{21}$$

给出．式中 λ_j 是 $\boldsymbol{A}$ 的相异特征值，d_j 是 $\boldsymbol{A}$ 的若当形 $\boldsymbol{J}$ 中包含 λ_j 的分块的最大阶数，即

$$d_j=\max_K n_{jk} \quad j=1,\ \cdots,\ \sigma \tag{22}$$

d_j 等于满足

$$\operatorname{rank}(\lambda_j\boldsymbol{I}-\boldsymbol{A})^{k+1}=\operatorname{rank}(\lambda_j\boldsymbol{I}-\boldsymbol{A})^k \tag{23}$$

的最小自然数 k.

［系 1］ 对于矩阵 $\boldsymbol{A}$，下列性质是等价的．

(i) 所有特征值的几何重复度是 1.

(ii) 特征多项式和最小多项式一致.

［系 2］ 对于矩阵 $\boldsymbol{A}$，下列性质是等价的.

(i) $\boldsymbol{A}$ 是单纯矩阵.

(ii) $d_j=1,\ j=1,\ \cdots,\ \sigma$

(问题 1) $\boldsymbol{A}^k=\boldsymbol{0}$ 且 $\boldsymbol{A}^j\neq\boldsymbol{0}(j=1,\ 2,\ \cdots,\ k-1)$，是否知道 $\boldsymbol{A}$ 的最小多项式及 $\boldsymbol{A}$ 的若当形是什么样的.

对于若当形用定理4很容易得出如下性质(不用若当形也可以直接证明该结果)[25, 60]．

［系 3］ 设 $\operatorname{adj}(\lambda\boldsymbol{I}-\boldsymbol{A})$ 全元素的最大公约式为 $d(\lambda)$，则

$$\psi(\lambda)=\psi_m(\lambda)d(\lambda) \tag{24}$$

成立.

[系 4]　对于多项式 f, g, 使 $f(\boldsymbol{A})=g(\boldsymbol{A})$ 的充分必要条件是[1]

$$f^{(l)}(\lambda_j)=g^{(l)}(\lambda_j) \quad j=1,\ 2,\ \cdots,\ \sigma;\ l=0,\ 1,\ \cdots,\ d_j-1 \tag{25}$$

(证明)　假定 $f(\boldsymbol{A})=g(\boldsymbol{A})$. 因 $h(\lambda)\triangleq f(\lambda)-g(\lambda)$ 是零化多项式, 根据定理 2, 则 h 能用最小多项式整除, 即对于某个多项式 q 可以写成 $h(\lambda)=q(\lambda)\psi_m(\lambda)$. 因此, (25) 式成立.

反之, 若 f 和 g 在 $\boldsymbol{A}$ 的谱上一致, 则多项式 h 能用 ψ_m 除尽. 因此, $h(\boldsymbol{A})=\boldsymbol{0}$, 即

$$f(\boldsymbol{A})=g(\boldsymbol{A})$$

证明完毕

在系 4 中若假定 $g(\lambda)=\psi_m(\lambda)$, 则下列结果成立.

[系 5]　对于某多项式 f, $f(\boldsymbol{A})=\boldsymbol{0}$ 的充分必要条件是, 在 $\boldsymbol{A}$ 的谱上 f 的值为零, 即 $f^{(l)}(\lambda_j)=0,\ j=1,\ 2,\ \cdots,\ \sigma;\ l=0,\ 1,\ \cdots,\ d_j-1$.

对于单纯矩阵, 矩阵多项式可按 (7) 式分解. 下面我们来讨论, 在包含非单纯矩阵的情况下, 将该性质一般化. 设所给矩阵 $\boldsymbol{A}$ 用

$$\boldsymbol{A}=\boldsymbol{TJT}^{-1} \tag{25}$$

变换成若当形. 根据矩阵多项式的性质 (ii), 因

$$f(\boldsymbol{A})=\boldsymbol{T}f(\boldsymbol{J})\boldsymbol{T}^{-1} \tag{26}$$

成立. 为了分解 $f(\boldsymbol{A})$, 则可以先分解 $f(\boldsymbol{J})$, 然后把它代入 (26) 式即可. 下面来讨论 $f(\boldsymbol{J})$ 的分解.

参看 B-I-11 中 (19) 式

$$\boldsymbol{J}=\sum_{j=1}^{\sigma}\hat{\boldsymbol{J}}_j \tag{27}$$

式中 $\hat{\boldsymbol{J}}_j$ 是 $\boldsymbol{J}$ 中的第 j 个子块, 除了这种子块以外, 其它子块均为 $\boldsymbol{0}$. 即

$$\hat{\boldsymbol{J}}_j\triangleq\begin{bmatrix}\boldsymbol{0} & & \boldsymbol{0}\\ & \boxed{\boldsymbol{J}_j} & \\ \boldsymbol{0} & & \boldsymbol{0}\end{bmatrix}\leftarrow \text{第 } j \text{ 个子块} \tag{28}$$

当然

$$f(\boldsymbol{J})=\sum_{j=1}^{\sigma}f(\hat{\boldsymbol{J}}_j) \tag{29}$$

$$f(\hat{\boldsymbol{J}}_j)=\begin{bmatrix}\boldsymbol{0} & & \boldsymbol{0}\\ & \boxed{f(\boldsymbol{J}_j)} & \\ \boldsymbol{0} & & \boldsymbol{0}\end{bmatrix}\leftarrow \text{第 } j \text{ 个子块} \tag{30}$$

而且, 由 B-I-11 中 (20) 式, 得

$$f(\boldsymbol{J}_j)=f(\boldsymbol{J}_{j1})\dotplus\cdots\dotplus f(\boldsymbol{J}_{j\alpha_j}) \tag{31}$$

因此, 必须先求 $f(\boldsymbol{J}_{jk})$. 由 B-I-11 中 (21) 式, $\boldsymbol{J}_{jk}$ 为

1) $f^{(l)}(\lambda_j)$ 表示 $[d^lf(\lambda)/d\lambda^l]_{\lambda=\lambda_j}$. (25) 式成立, 称为函数 f 和 g 在 $\boldsymbol{A}$ 的谱上 (即在 $\boldsymbol{A}$ 的特征值上——译者注) 一致. 而且, 按照 $f^{(l)}(\lambda_j)$, $j=1,\ 2,\ \cdots,\ \sigma;\ l=0,\ 1,\ \cdots,\ d_j-1$ 定义的 f, 称为在 $\boldsymbol{A}$ 的谱上定义的函数.

$$
\boldsymbol{J}_{jk}=\begin{bmatrix} \lambda_j & 1 & & & \\ & \lambda_j & 1 & \boldsymbol{0} & \\ & & \ddots & \ddots & \\ & \boldsymbol{0} & & \ddots & 1 \\ & & & & \lambda_j \end{bmatrix}=n_{jk}\times n_{jk} \tag{32}
$$

$\boldsymbol{J}_{jk}$ 的最小多项式与其特征多项式一致，得

$$
\psi_m(\lambda)=\psi(\lambda)=(\lambda-\lambda_j)^{n_{jk}}. \tag{33}
$$

因此，若利用(22)式的 d_j 定义多项式

$$
g(\lambda)\triangleq f(\lambda_j)+f^{(1)}(\lambda_j)(\lambda-\lambda_j)+\cdots+\frac{f^{(d_j-1)}(\lambda_j)}{(d_j-1)!}(\lambda-\lambda_j)^{d_j-1} \tag{34}
$$

则 $f(\lambda)$ 和 $g(\lambda)$ 在 $\boldsymbol{J}_{jk}(k=1,\ 2,\ \cdots,\ \alpha_j)$ 的谱上一致. 由系 4, 则

$$
\begin{aligned} f(\boldsymbol{J}_{jk})=g(\boldsymbol{J}_{jk})&=f(\lambda_j)\boldsymbol{I}_{n_{jk}}+f^{(1)}(\lambda_j)(\boldsymbol{J}_{jk}-\lambda_j\boldsymbol{I}_{n_{jk}})+\cdots \\ &+\frac{f^{(d_j-1)}(\lambda_j)}{(d_j-1)!}(\boldsymbol{J}_{jk}-\lambda_j\boldsymbol{I}_{n_{jk}})^{d_j-1}\quad k=1,\ 2,\ \cdots,\ \alpha_j \end{aligned} \tag{35}
$$

成立. 将(35)式代入(31)式，得

$$
\begin{aligned} f(\boldsymbol{J}_j)=g(\boldsymbol{J}_{j1})\dotplus\cdots\dotplus g(\boldsymbol{J}_{j\alpha_j})&=f(\lambda_j)\boldsymbol{I}_{mj}+f^{(1)}(\lambda_j)(\boldsymbol{J}_j-\lambda_j\boldsymbol{I}_{mj})+\cdots \\ &+\frac{f^{(d_j-1)}(\lambda_j)}{(d_j-1)!}(\boldsymbol{J}_j-\lambda_j\boldsymbol{I}_{mj})^{d_j-1}. \end{aligned} \tag{36}
$$

将(36)式代入(30)式，便得到求 $f(\hat{\boldsymbol{J}}_j)$ 的步骤如下:

1. 将 $\hat{\boldsymbol{J}}_j$ 的第 j 个子块 $\boldsymbol{J}_j$ 换成 $\boldsymbol{I}_{mj}$，设得到的为 $\boldsymbol{H}_j^0$;
2. 由 $\boldsymbol{H}_j^1\triangleq\hat{\boldsymbol{J}}_j-\lambda_j\boldsymbol{H}_j^0$ 求 $\boldsymbol{H}_j^1$;
3. 由 $\boldsymbol{H}_j^l\triangleq\dfrac{1}{l!}(\boldsymbol{H}_j^1)^l$ 求 $\boldsymbol{H}_j^l(l\geqslant2)$;
4. 由 $f(\hat{\boldsymbol{J}}_j)=\sum\limits_{l=0}^{d_j-1}\left(\dfrac{d^lf}{d\lambda^l}\right)_{\lambda=\lambda_j}\boldsymbol{H}_j^l$ 可以求出 $f(\hat{\boldsymbol{J}}_j)$.

将"4."中求得的 $f(\hat{\boldsymbol{J}}_j)$ 代入(29)式，然后再把得到的结果代入(26)式，得

$$
f(\boldsymbol{A})=\boldsymbol{T}\left\{\sum_{j=1}^{\sigma}\sum_{l=0}^{d_j-1}\left(\frac{d^lf}{d\lambda^l}\right)_{\lambda=\lambda_j}\boldsymbol{H}_j^l\right\}\boldsymbol{T}^{-1}=\sum_{j=1}^{\sigma}\sum_{l=0}^{d_j-1}\left(\frac{d^lf}{d\lambda^l}\right)_{\lambda=\lambda_j}\boldsymbol{T}\boldsymbol{H}_j^l\boldsymbol{T}^{-1} \tag{37}
$$

根据(37)式，对于所给的 $\boldsymbol{A}$，可以得到下列分解 $f(\boldsymbol{A})$ 的步骤:

0. 将所给 $\boldsymbol{A}$ 变换成若当形，设为 $\boldsymbol{A}=\boldsymbol{TJT}^{-1}$. 将若当形 $\boldsymbol{J}$ 按(27), (28)式分解.

根据上述步骤 1～3, 求出 $\boldsymbol{H}_j^l(l=0,\ 1,\ \cdots)$.

4.′ 由 $\boldsymbol{E}_j^l\triangleq\boldsymbol{TH}_j^l\boldsymbol{T}^{-1}$ 求 $\boldsymbol{E}_j^l$

5. $f(\boldsymbol{A})$ 可以由

$$
f(\boldsymbol{A})=\sum_{j=1}^{\sigma}\sum_{l=0}^{d_j-1}\left(\frac{d^lf}{d\lambda^l}\right)_{\lambda=\lambda_i}\boldsymbol{E}_j^l \tag{38}
$$

给出.

［例题］ 设 $\boldsymbol{A}=\boldsymbol{TJT}^{-1}$,

$$
\boldsymbol{J}=\left[\begin{array}{ccc|c|c|cc}
\lambda_1 & 1 & 0 & & & & \\
0 & \lambda_1 & 1 & & \boldsymbol{0} & & \\
0 & 0 & \lambda_1 & & & & \\
\hline
 & & & \lambda_1 & & & \\
\hline
 & & & & \lambda_1 & & \\
\hline
 & \boldsymbol{0} & & & & \lambda_2 & 1 \\
 & & & & & 0 & \lambda_2
\end{array}\right] \tag{39}
$$

则

$$
\boldsymbol{H}_1^0=\left[\begin{array}{ccccc|c}
1 & & & & & \\
 & 1 & & \boldsymbol{0} & & \\
 & & 1 & & & \boldsymbol{0} \\
\boldsymbol{0} & & & 1 & & \\
 & & & & 1 & \\
\hline
 & & \boldsymbol{0} & & & \boldsymbol{0}
\end{array}\right],
$$

$$
\boldsymbol{H}_2^0=\left[\begin{array}{c|cc}
\boldsymbol{0} & \multicolumn{2}{c}{\boldsymbol{0}} \\
\hline
\boldsymbol{0} & 1 & 0 \\
 & 0 & 1
\end{array}\right],
$$

$$
\boldsymbol{H}_1^1=\left[\begin{array}{ccc|c|c}
0 & 1 & 0 & & \\
0 & 0 & 1 & \boldsymbol{0} & \boldsymbol{0} \\
0 & 0 & 0 & & \\
\hline
 & \boldsymbol{0} & & \boldsymbol{0} & \\
\hline
 & \boldsymbol{0} & & & \boldsymbol{0}
\end{array}\right],
$$

$$
\boldsymbol{H}_2^1=\left[\begin{array}{c|cc}
\boldsymbol{0} & \multicolumn{2}{c}{\boldsymbol{0}} \\
\hline
\boldsymbol{0} & 0 & 1 \\
 & 0 & 0
\end{array}\right]
$$

$$
\boldsymbol{H}_1^2=\left[\begin{array}{ccc|c}
0 & 0 & 1 & \\
0 & 0 & 0 & \boldsymbol{0} \\
0 & 0 & 0 & \\
\hline
 & \boldsymbol{0} & & \boldsymbol{0}
\end{array}\right] \qquad \begin{array}{ll} \boldsymbol{H}_2^l=\boldsymbol{0} & l\geqslant 2 \\ \boldsymbol{H}_1^l=\boldsymbol{0} & l\geqslant 3 \end{array}
$$

(问题 2)　对于(38)式的 $\boldsymbol{E}_j^l$，试证明如下性质:

(i) $\boldsymbol{E}_{j_1}^{l_1}\boldsymbol{E}_{j_2}^{l_2}=\boldsymbol{0} \quad j_1\neq j_2 \quad {}^\forall l_1 l_2$

(ii) $\sum_{j=1}^{\sigma}\boldsymbol{E}_j^0=\boldsymbol{I}_n$

(iii) $\boldsymbol{E}_j^l=\dfrac{1}{l!}(\boldsymbol{A}-\lambda_j\boldsymbol{E}_j^0)^l \quad l\geqslant 1,\ j=1,\ \cdots,\ \sigma$

(注意)

不用若当形也可以求出 $\boldsymbol{E}_j^l$，该方法将在矩阵函数一章予以说明.

B-I-13 矩阵函数[25]

定义矩阵函数有几个途径．例如，在B-I-6中，与对于任意$\lambda\in C$绝对收敛的纯量幂级数

$$e^{\lambda}=\sum_{k=0}^{\infty}\frac{\lambda^k}{k!},\quad \cos\lambda=\sum_{k=0}^{\infty}(-1)^k\frac{\lambda^{2k}}{(2k)!},\quad \sin\lambda=\sum_{k=0}^{\infty}(-1)^k\frac{\lambda^{2k+1}}{(2k+1)!}$$

相对应，分别将矩阵函数e^{A}, $\sin \boldsymbol{A}$, $\cos \boldsymbol{A}$定义成下列绝对收敛的矩阵幂级数之和

$$e^{A}=\sum_{k=0}^{\infty}\frac{\boldsymbol{A}^k}{k!},\quad \cos\boldsymbol{A}=\sum_{k=0}^{\infty}(-1)^k\frac{\boldsymbol{A}^{2k}}{(2k)!},\quad \sin\boldsymbol{A}=\sum_{k=0}^{\infty}(-1)^k\frac{\boldsymbol{A}^{2k+1}}{(2k+1)!}$$

下面先从更普遍的观点给出矩阵函数的定义，它包含上述矩阵函数的定义，并证明由此得到的各种性质．当然，作为上述幂级数和的矩阵函数，在其幂级数收敛的情况下，和按此一般定义得到的矩阵函数一致．关于这一点以后再讲．

矩阵函数的定义

我们已经知道，当纯量多项式$f(\lambda)$用

$$f(\lambda)=a_0+a_1\lambda+\cdots+a_l\lambda^l,\ a_k\in C(k=0,\ 1,\ \cdots,\ l)$$

给出时，矩阵$\boldsymbol{A}\in C^{n\times n}$的多项式$f(\boldsymbol{A})$可以定义成

$$f(\boldsymbol{A})\triangleq a_0\boldsymbol{I}+a_1\boldsymbol{A}+\cdots+a_l\boldsymbol{A}^l.$$

在下面将要叙述的一般定义中，对于矩阵多项式成立的几个基本性质，对于更一般的函数f（的矩阵函数）也成立．

象以前，我们先设$\boldsymbol{A}\in C^{n\times n}$的最小多项式为

$$\psi_m(\lambda)=(\lambda-\lambda_1)^{d_1}(\lambda-\lambda_2)^{d_2}\cdots(\lambda-\lambda_s)^{d_s}.$$

式中λ_1, λ_2, $\cdots$, λ_s是$\boldsymbol{A}$的相异特征值．根据B-I-12中[系4]，对于在$\boldsymbol{A}$的谱上一致的两个纯量多项式f, g，其矩阵多项式完全一致．即，当且仅当

$$f^{(l)}(\lambda_k)=g^{(l)}(\lambda_k),\quad k=1,\ 2,\ \cdots,\ s;\ l=0,\ 1,\ \cdots,\ d_k-1$$

时，$f(\boldsymbol{A})=g(\boldsymbol{A})$．也就是说，将$\boldsymbol{A}$固定的情况下，任意矩阵多项式$f(\boldsymbol{A})$可仅由$\boldsymbol{A}$谱上的$f$值完全规定．我们将这样来定义矩阵函数，使得该性质对于包含多项式的一般函数也成立．亦即，对于在$\boldsymbol{A}$的谱上一致的（不只限于多项式）函数f, g，要使$f(\boldsymbol{A})=g(\boldsymbol{A})$成立．

特别是，若g选成（和f在$\boldsymbol{A}$的谱上一致的）多项式，因多项式矩阵$g(\boldsymbol{A})$已经定义，则$f(\boldsymbol{A})$亦可按$f(\boldsymbol{A})=g(\boldsymbol{A})$定义．

那么，为了能够象这样地定义$f(\boldsymbol{A})$，则f必须是在$\boldsymbol{A}$的谱上可以定义的函数，即$f^{(l)}(\lambda_k)$ $(k=1,\ 2,\ \cdots,\ s;\ l=0,\ 1,\ \cdots,\ d_k-1)$必须是有限值．关于这一点，若以在包含$\boldsymbol{A}$的相异特征值$\lambda_1$, λ_2, $\cdots$, λ_s的开域上可以定义的解析函数为对象，则实用上是足够了．以下，凡提到矩阵函数$f(\boldsymbol{A})$时，都不再加以说明，认为f满足这个条件．

［定义］“设和给定函数f在$\boldsymbol{A}$的谱上一致的任意多项式为g，则f的矩阵函数可由

$$f(\boldsymbol{A}) \triangleq g(\boldsymbol{A}) \tag{1}$$

定义.”

在 $\boldsymbol{A}$ 的谱上和 f 一致的任何一个多项式，都给出同样的 $f(\boldsymbol{A})$. 其原因是，若 g_1, g_2 是这样的多项式，由 B-I-12 中[系 4]，则 $g_1(\boldsymbol{A})=g_2(\boldsymbol{A})$. 因此，也可以用 $\boldsymbol{A}$ 的特征多项式代替 $\boldsymbol{A}$ 的最小多项式定义 $f(\boldsymbol{A})$. 现假定

$$\psi(\lambda) \triangleq \det[\lambda \boldsymbol{I}-\boldsymbol{A}] \triangleq (\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_s)^{m_s}$$

设 $\tilde{g}$ 为与 $\boldsymbol{A}$ 的特征多项式有关的，在 $\boldsymbol{A}$ 的谱上与 f 一致的多项式. 即，当

$$f^{(l)}(\lambda_k)=\tilde{g}^{(l)}(\lambda_k) \quad k=1,\ 2,\ \cdots,\ s;\ l=0,\ 1,\ \cdots,\ m_k-1 \tag{2a}$$

成立时，可以写成

$$f(\boldsymbol{A}) \triangleq \tilde{g}(\boldsymbol{A}) \tag{2b}$$

其原因是，因 g 和 $\tilde{g}$ 在 $\boldsymbol{A}$ 的谱上一致，则 $g(\boldsymbol{A})=\tilde{g}(\boldsymbol{A})$，其中每一个都给出同样的矩阵函数.

下面，我们来说明直接由上述定义得出的矩阵函数的几个基本性质.

矩阵函数的性质

（i） $\boldsymbol{A}$ 和 $f(\boldsymbol{A})$ 可换，即

$$\boldsymbol{A}f(\boldsymbol{A})=f(\boldsymbol{A})\boldsymbol{A} \tag{3}$$

（ii） 函数和(积)的矩阵函数，可以用矩阵函数的和(积)给出.

$$(f_1+f_2)(\boldsymbol{A})=f_1(\boldsymbol{A})+f_2(\boldsymbol{A}) \tag{4}$$

$$(f_1f_2)(\boldsymbol{A})=f_1(\boldsymbol{A})f_2(\boldsymbol{A}) \tag{5}$$

(iii) 若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 相似，则 $f(\boldsymbol{A})$ 和 $f(\boldsymbol{B})$ 也相似.

$$\boldsymbol{B}=\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}(\boldsymbol{T}\in C^{n\times n}) \Rightarrow f(\boldsymbol{B})=\boldsymbol{T}^{-1}f(\boldsymbol{A})\boldsymbol{T} \tag{6}$$

(iv) 若 $\boldsymbol{A}$ 是分块对角矩阵，则 $f(\boldsymbol{A})$ 亦为分块对角矩阵.

$$\boldsymbol{A}=\boldsymbol{A}_1 \dotplus \boldsymbol{A}_2 \dotplus \cdots \dotplus \boldsymbol{A}_\sigma \Rightarrow f(\boldsymbol{A})=f(\boldsymbol{A}_1) \dotplus f(\boldsymbol{A}_2) \dotplus \cdots \dotplus f(\boldsymbol{A}_\sigma) \tag{7}$$

(v) 设 $g(x_1,\ x_2)$ 是关于两个变量 x_1, x_2 的多项式. 当给出某函数 f_1, f_2 时，若在 $\boldsymbol{A}$ 的谱上

$$g(f_1(\lambda),\ f_2(\lambda))=0 \tag{8}$$

则矩阵恒等式

$$g(f_1(\boldsymbol{A}),\ f_2(\boldsymbol{A}))=\boldsymbol{0} \tag{9}$$

成立.

(证明)

(i), (ii): 若 f 为多项式，显然(3), (4), (5)式是正确的. 在 f 为解析函数的情况下，根据定义，因为 $f(\boldsymbol{A})$ 等于多项式矩阵 $g(\boldsymbol{A})$，则这些性质亦成立.

(iii): $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 有同样的最小多项式. 若在 $\boldsymbol{A}$(及 $\boldsymbol{B}$)的谱上和 f 一致的多项式为 g，则 $f(\boldsymbol{A})=g(\boldsymbol{A})$, $f(\boldsymbol{B})=g(\boldsymbol{B})$. 而且，因 $g(\boldsymbol{B})=\boldsymbol{T}^{-1}g(\boldsymbol{A})\boldsymbol{T}$(B-I-12 中多项式矩阵的性质(ii))，则 $f(\boldsymbol{B})=\boldsymbol{T}^{-1}f(\boldsymbol{A})\boldsymbol{T}$.

(iv): 设 $\boldsymbol{A}_i$ 的特征多项式为 ψ_i. 显然，$\boldsymbol{A}$ 的特征多项式 $\psi(\lambda)=\psi_1(\lambda)\psi_2(\lambda)\cdots\psi_\sigma(\lambda)$. 若令在特征多项式 ψ 的谱上和 f 一致的多项式为 g，根据关于(2)式的注意事项，因为 $f(\boldsymbol{A})=g(\boldsymbol{A})$, $f(\boldsymbol{A}_i)=g(\boldsymbol{A}_i)$, $i=1,\ 2,\ \cdots,\ \sigma$, $g(\boldsymbol{A})=g(\boldsymbol{A}_1) \dotplus g(\boldsymbol{A}_2)\cdots \dotplus g(\boldsymbol{A}_\sigma)$，则 $f(\boldsymbol{A})=g(\boldsymbol{A})=f(\boldsymbol{A}_1) \dotplus f(\boldsymbol{A}_2)\cdots \dotplus f(\boldsymbol{A}_\sigma)$.

(v)：设和 f_1, f_2 在 $\boldsymbol{A}$ 的谱上一致的多项式为 g_1, g_2，λ 的多项式 $g(g_1(\lambda), g_2(\lambda)) \triangleq q(\lambda)$.

在 $\boldsymbol{A}$ 的谱上 $g(f_1(\lambda), f_2(\lambda))=0$ 意味着，同样在 $\boldsymbol{A}$ 的谱上 $q(\lambda)=g(g_1(\lambda), g_2(\lambda))=0$. 因此，$q(\boldsymbol{A})=0$，即

$$g(g_1(\boldsymbol{A}), g_2(\boldsymbol{A}))=g(f_1(\boldsymbol{A}), f_2(\boldsymbol{A}))=\boldsymbol{0}.$$

[例 1] 考虑若当标准形矩阵

$$\boldsymbol{J}=\begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \lambda_i \end{bmatrix} \in C^{n\times n} \tag{10}$$

$\boldsymbol{J}$ 的最小多项式与特征多项式一致

$$\psi_m(\lambda)=\psi(\lambda)=(\lambda-\lambda_i)^n.$$

容易确认，在 $\boldsymbol{J}$ 的谱上与 f 一致的多项式 g(之一)可以由

$$g(\lambda)=f(\lambda_i)+f^{(1)}(\lambda_i)(\lambda-\lambda_i)+\cdots\frac{f^{(n-1)}(\lambda_i)}{(n-1)!}(\lambda-\lambda_i)^{n-1}$$

给出.

因此，根据定义(1)式，$\boldsymbol{J}$ 的矩阵函数

$$f(\boldsymbol{J})=\begin{bmatrix} f(\lambda_i) & f^{(1)}(\lambda_i) & \frac{1}{2!}f^{(2)}(\lambda_i) & \cdots & \frac{1}{(n-1)!}f^{(n-1)}(\lambda_i) \\ 0 & f(\lambda_i) & f^{(1)}(\lambda_i) & \cdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & f^{(1)}(\lambda_i) \\ 0 & 0 & \cdots & \cdots & f(\lambda_i) \end{bmatrix} \tag{11}$$

[例 2] 由性质(v)可见，大家所熟悉的几个关于纯量函数的恒等式，对于矩阵也成立.

$$\begin{aligned} &\cos^2\lambda+\sin^2\lambda=1 \Rightarrow \cos^2\boldsymbol{A}+\sin^2\boldsymbol{A}=\boldsymbol{I} \quad \forall \boldsymbol{A}\in C^{n\times n} \\ &e^{\lambda}e^{-\lambda}=1 \qquad\quad\ \Rightarrow e^{\boldsymbol{A}}\cdot e^{-\boldsymbol{A}}=\boldsymbol{I} \qquad\quad \forall \boldsymbol{A}\in C^{n\times n} \\ &e^{j\lambda}=\cos\lambda+j\sin\lambda \Rightarrow e^{j\boldsymbol{A}}=\cos\boldsymbol{A}+j\sin\boldsymbol{A}, \ \forall \boldsymbol{A}\in C^{n\times n} \end{aligned}$$

$f(\boldsymbol{A})$ 用若当标准形表示(标准形 1)

将矩阵函数的性质(iii)，(iv)和例 1 的结果结合起来，则利用矩阵 $\boldsymbol{A}$ 的若当标准形可以推导出矩阵函数的一个标准形.

设 $\boldsymbol{A}\in C^{n\times n}$ 的若当标准形 $\boldsymbol{J}=\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$(参看 B-I-11 中定理 6)为

$$\boldsymbol{J}=\begin{bmatrix} \boldsymbol{J}_1 & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{J}_2 & \cdots & \boldsymbol{0} \\ \vdots & & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{J}_s \end{bmatrix} = n\times n \tag{12a}$$

$$\boldsymbol{J}_i=\begin{bmatrix}\boldsymbol{J}_{i1} & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \boldsymbol{J}_{i2} & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{J}_{i\alpha_i}\end{bmatrix}=m_i\times m_i,\quad \sum_{i=1}^{s}m_i=n\quad i=1,\ 2,\ \cdots,\ s \tag{12b}$$

$$\boldsymbol{J}_{ij}=\begin{bmatrix}\lambda_i & 1 & \cdots & 0\\ 0 & \lambda_i & \ddots & \vdots\\ \vdots & \vdots & \ddots & 1\\ 0 & 0 & \cdots & \lambda_i\end{bmatrix}=n_{ij}\times n_{ij},\quad \sum_{j=1}^{\alpha_i}n_{ij}=m_i\quad j=1,\ 2,\ \cdots,\ \alpha_i \tag{12c}$$

则 $f(\boldsymbol{A})$ 可以表示如下.

$$f(\boldsymbol{A})=\boldsymbol{T}f(\boldsymbol{J})\boldsymbol{T}^{-1} \tag{13a}$$

$$f(\boldsymbol{J})=\begin{bmatrix}f(\boldsymbol{J}_1) & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & f(\boldsymbol{J}_2) & & \mathbf{0}\\ \vdots & & \ddots & \vdots\\ \mathbf{0} & \cdots & \cdots & f(\boldsymbol{J}_s)\end{bmatrix}=n\times n \tag{13b}$$

$$f(\boldsymbol{J}_i)=\begin{bmatrix}f(\boldsymbol{J}_{i1}) & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & f(\boldsymbol{J}_{i2}) & \cdots & \mathbf{0}\\ \vdots & & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & f(\boldsymbol{J}_{i\alpha_i})\end{bmatrix}=m_i\times m_i\quad i=1,\ 2,\ \cdots,\ s \tag{13c}$$

$$f(\boldsymbol{J}_{ij})=\begin{bmatrix}f(\lambda_i) & f^{(1)}(\lambda_i) & \cdots & \dfrac{f^{(n_{ij}-1)}(\lambda_i)}{(n_{ij}-1)!}\\ & \ddots & \ddots & \vdots\\ & \mathbf{0} & \ddots & f^{(1)}(\lambda_i)\\ & & & f(\lambda_i)\end{bmatrix}=n_{ij}\times n_{ij}\quad j=1,\ 2,\ \cdots,\ \alpha_i \tag{13d}$$

为了与 $f(\boldsymbol{A})$ 的其它表示方法相区别，称上述表示为标准形 1.

特别是，在最小多项式为 $\psi_m(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_s)$ 的情况下(即，$d_1=d_2=\cdots=d_s=1$，也包含 $\boldsymbol{A}$ 有 n 个相异特征值的情况)，$\boldsymbol{A}$ 的若当标准形变成对角矩阵.

因

$$\boldsymbol{J}=\begin{bmatrix}\boldsymbol{J}_1 & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \boldsymbol{J}_2 & \cdots & \mathbf{0}\\ \vdots & \vdots & & \vdots\\ \mathbf{0} & \mathbf{0} & & \boldsymbol{J}_s\end{bmatrix}=n\times n \tag{14a}$$

$$\boldsymbol{J}_i=\begin{bmatrix}\lambda_i & 0 & \cdots & 0\\ 0 & \lambda_i & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_i\end{bmatrix}=m_i\times m_i,\quad \sum_{i=1}^{s}m_i=n\quad i=1,\ 2,\ \cdots,\ s \tag{14b}$$

则 $f(\boldsymbol{A})$ 的标准形 1 为

$$f(\boldsymbol{A})=\boldsymbol{T}f(\boldsymbol{J})\boldsymbol{T}^{-1} \tag{15a}$$

$$f(\boldsymbol{J})=\begin{bmatrix} f(\boldsymbol{J}_1) & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & f(\boldsymbol{J}_2) & \cdots & \boldsymbol{0} \\ \vdots & & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & f(\boldsymbol{J}_s) \end{bmatrix}=n\times n \tag{15b}$$

$$f(\boldsymbol{J}_i)=\begin{bmatrix} f(\lambda_i) & 0 & \cdots & 0 \\ 0 & f(\lambda_i) & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & f(\lambda_i) \end{bmatrix}=m_i\times m_i,\ \sum_{i=1}^{s} m_i=n \quad i=1,\ 2,\ \cdots,\ s \tag{15c}$$

注意，$f(\boldsymbol{J})$ 也变成对角矩阵.

［例 3］ 当 $\boldsymbol{A}\in C^{n\times n}$ 的若当标准形 $\boldsymbol{J}=\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$ 按(12)式给出时，我们用标准形 1 来求矩阵函数 e^{At}.

因 $f(\lambda)=e^{\lambda t}$，则对于任意 $\boldsymbol{A}\in C^{n\times n}$ 和 $t\in R$，e^{At} 可定义，由(13)式得

$$e^{At}=\boldsymbol{T}e^{Jt}\boldsymbol{T}^{-1} \tag{16a}$$

$$e^{Jt}=\begin{bmatrix} e^{J_1t} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & e^{J_2t} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & e^{J_st} \end{bmatrix} \tag{16b}$$

$$e^{J_it}=\begin{bmatrix} e^{J_{i1}t} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & e^{J_{i2}t} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & & e^{J_{in_i}t} \end{bmatrix} \tag{16c}$$

$$e^{J_{ij}t}=\begin{bmatrix} e^{\lambda_it} & te^{\lambda_it} & \cdots & \dfrac{t(n_{ij}-1)}{(n_{ij}-1)!}e^{\lambda_it} \\ 0 & e^{\lambda_it} & \ddots & \vdots \\ \vdots & \vdots & \ddots & te^{\lambda_it} \\ 0 & 0 & \cdots\cdots & e^{\lambda_it} \end{bmatrix} \tag{16d}$$

$f(\boldsymbol{A})$用拉格朗日-西勒维斯特内插多项式表示（标准形 2）

根据矩阵函数的定义(1)式，可以利用和纯量函数 $f(\lambda)$ 在 $\boldsymbol{A}$ 的谱上一致的多项式 $g(\lambda)$ 确定 $f(\boldsymbol{A})$，$f(\boldsymbol{A})\triangleq g(\boldsymbol{A})$. 一般，这种多项式不是唯一的，而有无数个. 但是，其中比 $\boldsymbol{A}$ 的最小多项式次数小的只有所谓的拉格朗日-西勒维斯特内插多项式 (Lagrange-Sylvester interpolation polynominal).

设 $\boldsymbol{A}$ 的最小多项式为

$$\psi_m(\lambda)=(\lambda-\lambda_1)^{d_1}(\lambda-\lambda_2)^{d_2}\cdots(\lambda-\lambda_s)^{d_s} \tag{17}$$

$$\deg(\psi_m)\triangleq m=d_1+d_2+\cdots+d_s.$$

拉格朗日-西勒维斯特内插多项式 p 可求出如下. 假定 $\deg(p)<\deg(\psi_m)$，试将 p/ψ_m 展成部分分式

$$\frac{p(\lambda)}{\psi_m(\lambda)}=\sum_{k=1}^{s}\left[\frac{\alpha_{k1}}{(\lambda-\lambda_k)^{d_k}}+\frac{\alpha_{k2}}{(\lambda-\lambda_k)^{d_k-1}}+\cdots+\frac{\alpha_{kd_k}}{(\lambda-\lambda_k)}\right] \tag{18}$$

式中常数 $\alpha_{kl}(k=1, 2, \cdots, s, l=1, 2, \cdots, d_k)$ 与拉普拉斯变换有关，可以用大家所熟悉的公式

$$\alpha_{kl}=\frac{1}{(l-1)!}\left[\frac{d^{l-1}}{d\lambda^{l-1}}\left(\frac{(\lambda-\lambda_k)^{d_k}p(\lambda)}{\psi_m(\lambda)}\right)\right]^*_{\lambda=\lambda_k} \tag{19}$$

给出，而确定(19)式右边的数值只需要谱上的 p 值．而且根据假定，因 p 在谱上与 f 一致，则在(19)式中可将 $p(\lambda)$ 换成 $f(\lambda)$，最后 α_{kl} 可由

$$\alpha_{kl}=\frac{1}{(l-1)!}\left[\frac{d^{l-1}}{d\lambda^{l-1}}\left(\frac{(\lambda-\lambda_k)^{d_k}f(\lambda)}{\psi_m(\lambda)}\right)\right]_{\lambda=\lambda_k} \tag{20}$$

决定．知道 α_{kl} 后，所要求的 $p(\lambda)$ 可由(18)式得到

$$p(\lambda)=\sum_{k=1}^{s}[\alpha_{k1}+\alpha_{k2}(\lambda-\lambda_k)+\cdots+\alpha_{kd_k}(\lambda-\lambda_k)^{d_k-1}]\varphi_k(\lambda) \tag{21}$$

式中

$$\varphi_k(\lambda)\triangleq\frac{\psi_m(\lambda)}{(\lambda-\lambda_k)^{d_k}}=(\lambda-\lambda)^{d_1}\cdots(\lambda-\lambda_{k-1})^{d_{k-1}}(\lambda-\lambda_{k+1})^{d_{k+1}}\cdots(\lambda-\lambda_s)^{d_s}$$

因 $f(\boldsymbol{A})=p(\boldsymbol{A})$，则由此可求出矩阵函数 $f(\boldsymbol{A})$

$$f(\boldsymbol{A})=\sum_{k=1}^{s}[\alpha_{k1}\boldsymbol{I}+\alpha_{k2}(\boldsymbol{A}-\lambda_k\boldsymbol{I})+\cdots+\alpha_{kd_k}(\boldsymbol{A}-\lambda_k\boldsymbol{I})^{d_k-1}]\varphi_k(\boldsymbol{A}) \tag{22}$$

这个用拉格朗日-西勒维斯特内插多项式表示的 $f(\boldsymbol{A})$，我们称作标准形 2．在标准形 2 中，f 函数形状的影响出现在系数 α_{kl} 中(参看(20)式)，其它部分由 $\boldsymbol{A}$ 决定，与 f 无关．

作为特殊情况，设 $d_1=d_2=\cdots=d_s=1$，则

$$\alpha_{k1}=\frac{f(\lambda_k)}{\varphi_k(\lambda_k)},\ k=1, 2, \cdots, s \tag{23}$$

$$\varphi_k(\lambda)=(\lambda-\lambda_1)\cdots(\lambda-\lambda_{k-1})(\lambda-\lambda_{k+1})\cdots(\lambda-\lambda_s) \tag{24}$$

因此

$$f(\boldsymbol{A})=\sum_{k=1}^{s}\frac{f(\lambda_k)(\boldsymbol{A}-\lambda_1\boldsymbol{I})\cdots(\boldsymbol{A}-\lambda_{k-1}\boldsymbol{I})(\boldsymbol{A}-\lambda_{k+1}\boldsymbol{I})\cdots(\boldsymbol{A}-\lambda_s\boldsymbol{I})}{(\lambda_k-\lambda_1)\cdots(\lambda_k-\lambda_{k-1})(\lambda_k-\lambda_{k+1})\cdots(\lambda_k-\lambda_s)}. \tag{25}$$

该式也包括 $\boldsymbol{A}$ 有 n 个相异特征值的情况，这时在(25)式中令 $n=s$ 即可．

(问题 1) 试证明(21)式的 p 和 f 在 $\boldsymbol{A}$ 的谱上一致．

$f(\boldsymbol{A})$ 用矩阵分量表示(矩阵函数的基本公式，标准形 3)

将用拉格朗日-西勒维斯特内插多项式表示 $f(\boldsymbol{A})$ 的(25)式再稍整理一下，把与函数 f 有关的部分及由 $\boldsymbol{A}$ 的结构决定的部分清楚地分开表示出来．于是，便得到在系统理论应用上最方便的 $f(\boldsymbol{A})$ 的表示形式，称之为标准形 3，也有人称其为矩阵函数的基本公式 (fundamental formula for $f(\boldsymbol{A})$)[25, 28]．

首先，将(20)式代入拉格朗日-西勒维斯特内插多项式(21)式系数 α_{kl}，把含有 $f(\lambda_k)$，$f^{(1)}(\lambda_k)$，$\cdots$，$f^{(d_k-1)}(\lambda_k)$ $(k=1, 2, \cdots, s)$ 的项各自分开，整理后可以写成

$$p(\lambda)=\sum_{k=1}^{s}[f(\lambda_k)p_{k0}(\lambda)+f^{(1)}(\lambda_k)p_{k1}(\lambda)+\cdots+f^{(d_k-1)}(\lambda_k)p_{kd_k-1}(\lambda)] \tag{26}$$

* 原文误印成 $\alpha_{kl}=\frac{1}{(l-1)!}\left[\frac{d^{-1}}{d\lambda^{-1}}\left(\frac{(\lambda-\lambda_k)^{d_k}p(\lambda)}{\psi_m(\lambda)}\right)\right]_{\lambda=\lambda_k}$．——译者注

式中 $p_{kl}(k=1, 2, \cdots, s;\ l=0, 1, \cdots, d_k-1)$ 是与 λ 有关的多项式，其次数都小于 m. 实际上，直接将(20)式代入(21)式确定 $p_{kl}(\lambda)$ 是非常麻烦的，但若考虑到下述关系，也可以求出它的一般形.

在(26)式中，p_{kl} 是 $f^{(l)}(\lambda_k)$ 的系数多项式. 而且，因 p 是拉格朗日-西勒维斯特内插多项式，为使 p 和 f 在 $\boldsymbol{A}$ 的谱上一致，则下列关系

$$\left.\begin{aligned} p_{kl}^{(l)}(\lambda_k)&=1\\ p_{kl}^{(i)}(\lambda_k)&=0,\ i\neq l\\ p_{kl}^{(i)}(\lambda_j)&=0,\ j\neq k\end{aligned}\right\}\tag{27}$$

必须成立. 换句话说，f 是具有

$$\begin{aligned} f^{(l)}(\lambda_k)&=1\\ f^{(i)}(\lambda_k)&=0,\ i\neq l\\ f^{(i)}(\lambda_j)&=0,\ j\neq k\end{aligned}$$

性质的函数，p_{kl} 可以看成是和 f 在 $\boldsymbol{A}$ 的谱上一致的拉格朗日-西勒维斯特内插多项式. 因此 $p_{kl}(\lambda)$ 可求出如下

$$p_{kl}(\lambda)=\frac{1}{l!}(\lambda-\lambda_k)^l\phi_k(\lambda)\tag{28}$$

式中

$$\phi_k(\lambda)\triangleq n_k(\lambda)\varphi_k(\lambda)=n_k(\lambda)\frac{\psi_m(\lambda)}{(\lambda-\lambda_k)^{d_k}}\tag{29}$$

$n_k(\lambda)$ 是将 $1/\psi_m(\lambda)$ 展成部分分式时的分子多项式

$$\frac{1}{\psi_m(\lambda)}=\sum_{k=1}^{s}\frac{n_k(\lambda)}{(\lambda-\lambda_k)^{d_k}}.\tag{30}$$

于是，当 $p_{kl}(\lambda)$ 确定之后，便可以由(26)式得出 $f(\boldsymbol{A})$ 的标准形 3

$$f(\boldsymbol{A})=\sum_{k=1}^{s}[f(\lambda_k)p_{k0}(\boldsymbol{A})+f^{(1)}(\lambda_k)p_{k1}(\boldsymbol{A})+\cdots+f^{(d_k-1)}(\lambda_k)p_{kd_k-1}(\boldsymbol{A})]\tag{31}$$

式中 $p_{kl}(\boldsymbol{A})\ (k=1, 2, \cdots, s;\ l=0, 1, \cdots, d_k-1)$ 与 f 完全无关，是仅由 $\boldsymbol{A}$ 决定的常数矩阵，称为矩阵 $\boldsymbol{A}$ 的分量(component)或成分矩阵(constituent matrix).

特别是，在 $d_1=d_2=\cdots=d_s=1$ 的情况下，标准形 3 和(25)式一致. 即 $\boldsymbol{A}$ 的分量共有 s 个

$$p_{k0}(\boldsymbol{A})=\frac{(\boldsymbol{A}-\lambda_1\boldsymbol{I})\cdots(\boldsymbol{A}-\lambda_{k-1}\boldsymbol{I})(\boldsymbol{A}-\lambda_{k+1}\boldsymbol{I})\cdots(\boldsymbol{A}-\lambda_s\boldsymbol{I})}{(\lambda_k-\lambda_1)\cdots(\lambda_k-\lambda_{k-1})(\lambda_k-\lambda_{k+1})\cdots(\lambda_k-\lambda_s)}$$

$$k=1, 2, \cdots, s\tag{32}$$

(问题 2) 试证明(29)式 ϕ_k 满足 $\phi_k(\lambda_k)=1$, $\phi^{(l)}(\lambda_k)=0\ (l=1, 2, \cdots, d_k-1)$, $\phi_k^{(l)}(\lambda_j)=0\ (j\neq k)$. 利用这些关系试证明，(28)式 p_{kl} 具有(27)式的性质及(26)式 p 是拉格朗日-西勒维斯特内插多项式.

［例 4］ 设某矩阵 $\boldsymbol{A}\in C^{n\times n}$ 的最小多项式为

$$\psi_m(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)^2$$

现求其标准形 3.

$$\frac{1}{\psi_m(\lambda)}=\frac{1}{(\lambda_1-\lambda_2)^2(\lambda-\lambda_1)}-\frac{(\lambda-2\lambda_2+\lambda_1)}{(\lambda_1-\lambda_2)^2(\lambda-\lambda_2)^2}$$

$$\phi_1(\lambda)=\frac{(\lambda-\lambda_2)^2}{(\lambda_1-\lambda_2)^2},$$

$$\phi_2(\lambda)=\frac{(-\lambda-\lambda_1+2\lambda_2)(\lambda-\lambda_1)}{(\lambda_1-\lambda_2)^3}$$

$$p_{10}(\lambda)=\phi_1(\lambda),\ p_{20}(\lambda)=\phi_2(\lambda),\ p_{21}(\lambda)=(\lambda-\lambda_2)\phi_2(\lambda)$$

$$p(\lambda)=f(\lambda_1)p_{10}(\lambda)+f(\lambda_2)p_{20}(\lambda)+f^{(1)}(\lambda_2)p_{21}(\lambda)$$

$$=f(\lambda_1)\frac{(\lambda-\lambda_2)^2}{(\lambda_1-\lambda_2)^2}+f(\lambda_2)\frac{(-\lambda-\lambda_1+2\lambda_2)(\lambda-\lambda_1)}{(\lambda_1-\lambda_2)^2}$$

$$+f^{(1)}(\lambda_1)\frac{(-\lambda-\lambda_1-2\lambda_2)(\lambda-\lambda_1)(\lambda-\lambda_2)}{(\lambda_1-\lambda_2)^3}$$

$$f(\boldsymbol{A})=p(\boldsymbol{A})$$

［例5］ 现在来求 e^{At} 的标准形3．因 $f(\lambda)=e^{\lambda t}$，$f^{(l)}=t^l e^{\lambda t}$（注意，是对 λ 微分），则(31)式变成

$$e^{At}=\sum_{k=1}^{s}[p_{k0}(\boldsymbol{A})+tp_{k1}(\boldsymbol{A})+\cdots+t^{d_k-1}p_{kd_k-1}(\boldsymbol{A})]e^{\lambda_k t} \tag{33}$$

矩阵分量的性质

$\boldsymbol{A}$ 的分量 $p_{kl}(\boldsymbol{A})$ $(k=1,\ 2,\ \cdots,\ s;\ l=0,\ 1,\ \cdots,\ d_{k-1})$ 是线性独立的（当看成是复数体 C 上向量空间 $C^{n\times n}$ 中的向量时），使用 $f(\boldsymbol{A})$ 的标准形3（(31)式）时常常利用这个重要而且方便的性质．现说明如下．

［定理1］ 任意矩阵 $\boldsymbol{A}\in C^{n\times n}$ 的分量 $p_{kl}(\boldsymbol{A})$ $(k=1,\ 2,\ \cdots,\ s;\ l=0,\ 1,\ \cdots,\ d_k-1)$ 都是线性独立的．特别是，无论那个分量都不是零矩阵．

（证明） 我们先来证明纯量多项式 p_{kl} $(k=1,\ 2,\ \cdots,\ s;\ l=0,\ 1,\ \cdots,\ d_k-1)$ 是线性独立的．

现假定对于某个常数 $c_{kl}\in C$，

$$\sum_{k=1}^{s}\sum_{l=1}^{d_k-1}c_{kl}p_{kl}(\lambda)=0,\ \forall\lambda\in C \tag{34}$$

成立．这意味着（参看(26)式），满足

$$f^{(l)}(\lambda_k)=c_{kl}\quad(k=1,\ 2,\ \cdots,\ s;\ l=0,\ 1,\ \cdots,\ d_k-1)$$

的函数 f 的拉格朗日-西勒维斯特内插多项式 p 恒等于零，即 $p(\lambda)\equiv0$．因此，必须 $p^{(l)}(\lambda_k)=f^{(l)}(\lambda_k)=c_{kl}=0$，即 p_{kl} 是线性独立的．

但是，$p(\lambda)=\sum\limits_{k=1}^{s}\sum\limits_{l=0}^{d_k-1}c_{kl}p_{kl}(\lambda)$ 的次数小于最小多项式的次数．如上所述，因 p_{kl} 是线性独立的．则除了 $c_{kl}=0$ $(k=1,\ 2,\ \cdots,\ s;\ l=0,\ 1,\ \cdots,\ d_k-1)$ 之外，不会得到

$$p(\boldsymbol{A})=\sum_{k=1}^{s}\sum_{l=0}^{d_k-1}c_{kl}p_{kl}(\boldsymbol{A})=\boldsymbol{0} \tag{35}$$

即 $p_{kl}(\boldsymbol{A})$ 是线性独立的．

证明完毕

下面，再对求 $\boldsymbol{A}$ 的分量的方法作一些补充．

首先，在 $\boldsymbol{A}$ 的若当标准形 $\boldsymbol{J}=\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$ 求出的情况下，B-I-12 中已经证明，利用该标准形可以将任意多项式矩阵 $f(\boldsymbol{A})$ 表示成（B-I-12 中(38)* 式）

* 原文误为(47)式．——译者注

$$f(\boldsymbol{A})=\sum_{k=1}^{s}\sum_{l=0}^{d_k-1}f^{(l)}(\lambda_k)\boldsymbol{E}_k^l \tag{36}$$

同时给出求常数矩阵 $\boldsymbol{E}_k^l$ 的步骤. 因为一般的矩阵函数 $f(\boldsymbol{A})$ 都可以用和 f 在 $\boldsymbol{A}$ 的谱上一致的多项式给出,所以既使 f 不是多项式, (36)式也照样适用. 因为对于任意 f (36)式的右边必须与(31)式的右边一致,显然

$$p_{kl}(\boldsymbol{A})=\boldsymbol{E}_k^l \tag{37}$$

即可以用 $\boldsymbol{A}$ 的若当标准形确定 $\boldsymbol{A}$ 的分量.

我们知道[25], 对于 $\boldsymbol{A}$ 的分量计算, 利用 $[\lambda\boldsymbol{I}-\boldsymbol{A}]$ 的伴随矩阵是比较实用的方法. 因 $p_{kl}(\boldsymbol{A})$ 与 f 无关,仅决定于 $\boldsymbol{A}$, 则确定 $p_{kl}(\boldsymbol{A})$ 时可以选择最方便的函数. 现在考虑

$$f(z)=\frac{1}{\lambda-z} \tag{38}$$

这样的函数,则

$$f(\boldsymbol{A})=[\lambda\boldsymbol{I}-\boldsymbol{A}]^{-1}=\frac{\operatorname{adj}[\lambda\boldsymbol{I}-\boldsymbol{A}]}{\psi(\lambda)} \tag{39}$$

由 B-I-12 中[系 3], (39)式又可写成

$$\frac{\operatorname{adj}[\lambda\boldsymbol{I}-\boldsymbol{A}]}{\psi(\lambda)}=\frac{\boldsymbol{R}(\lambda)}{\psi_m(\lambda)}. \tag{40}$$

式中 ψ_m 是 $\boldsymbol{A}$ 的最小多项式, $\boldsymbol{R}(\lambda)$ 是以 λ 的多项式(都具有比 $\deg(\psi_m)$ 低的次数)为元素的矩阵. 将(40)式右边展成部分分式

$$\frac{\boldsymbol{R}(\lambda)}{\psi_m(\lambda)}=\sum_{k=1}^{s}\left[\frac{\boldsymbol{R}_{k0}}{\lambda-\lambda_k}+\frac{\boldsymbol{R}_{k1}}{(\lambda-\lambda_k)^2}+\cdots+\frac{\boldsymbol{R}_{kd_k-1}}{(\lambda-\lambda_k)^{d_k}}\right] \tag{41}$$

式中 $\boldsymbol{R}_{kl}$ 是常数矩阵,可以用和纯量部分分式展开公式(19)相同的形式

$$\boldsymbol{R}_{kl}=\frac{1}{(d_k-l-1)!}\left[\frac{d^{(d_k-l-1)}}{d\lambda^{(d_k-l-1)}}\left(\frac{(\lambda-\lambda_k)^{d_k}\boldsymbol{R}(\lambda)}{\psi_m(\lambda)}\right)\right]_{\lambda=\lambda_k} \tag{42}$$

给出. 此外,对于 $f(z)=1/(\lambda-z)$, 因

$$f^{(l)}(\lambda_k)=\frac{l!}{(\lambda-\lambda_k)^l}$$

则 $f(\boldsymbol{A})$ 的标准形 3 为

$$f(\boldsymbol{A})=\sum_{k=1}^{s}\left[\frac{1}{(\lambda-\lambda_k)}p_{k0}(\boldsymbol{A})+\frac{1}{(\lambda-\lambda_k)^2}p_{k1}(\boldsymbol{A})+\cdots+\frac{(d_{k-1})!}{(\lambda-\lambda_k)^{d_k}}p_{kd_k-1}(\boldsymbol{A})\right] \tag{43}$$

由(41), (43)式可以给出 $\boldsymbol{A}$ 的分量

$$p_{kl}(\boldsymbol{A})=\frac{1}{l!}\boldsymbol{R}_{kl},\ k=1,\ 2,\ \cdots,\ s;\ l=0,\ 1,\ 2,\ \cdots,\ d_k-1$$

$\boldsymbol{R}_{kl}$ 可以用(42)式计算.

$f(\boldsymbol{A}t)$ 对 t 的微分

上面讲过矩阵函数的定义,以及由此得到的几个性质(i)～(v). 它们表明,对于纯量函数成立的代数性质,在矩阵函数的情况下也都成立. 这里我们将要讨论,关于矩阵的微分是否可以用纯量情况下同样的公式. 即,对应于 $\frac{d}{dt}e^{\lambda t}=\lambda e^{\lambda t}$, 是否可以写成 $\frac{d}{dt}e^{\boldsymbol{A}t}=\boldsymbol{A}e^{\boldsymbol{A}t}$ 等问题. 关于这些问题,下面结果成立.

[定理 2] 设 $f(\lambda)$ 是整个复平面上的解析函数，$f(\lambda t)$ 对 $t(t\in \boldsymbol{R})$ 的微分为 $g(\lambda, t)$，即

$$\frac{\partial}{\partial t} f(\lambda t)=g(\lambda, t).$$

这时

$$\frac{\partial}{\partial t} f(\boldsymbol{A}t)=g(\boldsymbol{A}, t)$$

(证明) $g(\boldsymbol{A}, t)$ 的标准形 3 可以表示成

$$g(\boldsymbol{A}, t)=\sum_{k=1}^{s}\left[g(\lambda_k, t)p_{k0}(\boldsymbol{A})+\cdots+g^{(d_k-1)}(\lambda_k, t)p_{kd_k-1}(\boldsymbol{A})\right] \tag{44}$$

因

$$g^{(l)}(\lambda_k, t)\triangleq\left[\frac{\partial^l}{\partial\lambda^l}\frac{\partial}{\partial t}f(\lambda t)\right]_{\lambda=\lambda_k}=\frac{\partial}{\partial t}\left[\frac{\partial^l}{\partial\lambda^l}f(\lambda t)\right]_{\lambda=\lambda_k}$$

则 (44) 式变成

$$g(\boldsymbol{A}, t)=\frac{\partial}{\partial t}\sum_{k=1}^{s}\left[f(\lambda_k t)p_{k0}(\boldsymbol{A})+\cdots+f^{(d_k-1)}(\lambda_k t)p_{kd_k-1}(\boldsymbol{A})\right]$$

最后结果不外乎是 $\dfrac{\partial}{\partial t} f(\boldsymbol{A}t)$.

[例 6] 根据定理 2, 下列微分公式成立.

$$\frac{d}{dt}e^{\boldsymbol{A}t}=\boldsymbol{A}e^{\boldsymbol{A}t}=e^{\boldsymbol{A}t}\boldsymbol{A}$$

$$\frac{d}{dt}\cos\boldsymbol{A}t=-\boldsymbol{A}(\sin\boldsymbol{A}t)=-(\sin\boldsymbol{A}t)\boldsymbol{A}$$

$$\frac{d}{dt}\sin\boldsymbol{A}t=\boldsymbol{A}(\cos\boldsymbol{A}t)=(\cos\boldsymbol{A}t)\boldsymbol{A}$$

矩阵函数用幂级数表示

这里回到本章开始提出的问题，考虑作为绝对收敛幂级数和定义的矩阵函数和更一般的 (1) 式定义的矩阵函数之间的关系. 例如，在 B-I-6 中 $e^{\boldsymbol{A}t}$ 作为无限级数

$$\boldsymbol{I}+\boldsymbol{A}t+\frac{(\boldsymbol{A}t)^2}{2!}+\cdots+\frac{(\boldsymbol{A}t)^k}{k!}+\cdots$$

的和引入，另外根据 (1) 式定义，也可以用标准形 1((16) 式) 或标准形 3((33) 式) 给出. 根据下面定理，因为这几种不同方式可以给出完全相同的矩阵函数，不会出现不妥当之处，所以可以根据不同的目的选用.

[定理 3] 设 $\boldsymbol{A}$ 的最小多项式为 $\psi_m(\lambda)=(\lambda-\lambda_1)^{d_1}(\lambda-\lambda_2)^{d_2}\cdots(\lambda-\lambda_s)^{d_s}$, f 及 f_n $(n=1, 2, \cdots)$ 都是在包含 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 的开域上解析的函数序列.

当 $n\to\infty$, $f_n(\lambda)$ 在 $\boldsymbol{A}$ 的谱上收敛于 $f(\lambda)$ 时，即

$$\lim f_n^{(l)}(\lambda_k)=f^{(l)}(\lambda_k)\quad(k=1, 2, \cdots, s;\ l=0, 1, \cdots, d_k-1) \tag{45}$$

时，

$$\lim_{n\to\infty} f_n(\boldsymbol{A})=f(\boldsymbol{A}). \tag{46}$$

反之，若 (46) 式成立，则 (45) 式也成立.

（证明） 设(45)式成立．$f_n(\boldsymbol{A})$的标准形 3 可以表示成

$$f_n(\boldsymbol{A})=\sum_{k=1}^{s}[f_n(\lambda_k)p_{k0}(\boldsymbol{A})+f_n^{(1)}(\lambda_k)p_{k1}(\boldsymbol{A})+\cdots+f_n^{(d_k-1)}(\lambda_k)p_{kd_k-1}(\boldsymbol{A})] \tag{47}$$

$f(\boldsymbol{A})$ 的标准形 3 可以由

$$f(\boldsymbol{A})=\sum_{k=1}^{s}[f(\lambda_k)p_{k0}(\boldsymbol{A})+f^{(1)}(\lambda_k)p_{k1}(\boldsymbol{A})+\cdots+f^{(d_k-1)}(\lambda_k)p_{kd_k-1}(\boldsymbol{A})] \tag{48}$$

给出．显然，若(45)式成立，则 $\lim\limits_{n\to\infty}f_n(\boldsymbol{A})$存在，而且等于 $f(\boldsymbol{A})$.

反之，假定(46)式成立，根据[定理 1]，因 $\boldsymbol{A}$ 的分量是线性独立的，比较(47)，(48)式右边，则 $\lim\limits_{n\to\infty}f_n(\boldsymbol{A})-f(\boldsymbol{A})=\boldsymbol{0}$ 意味着 $\boldsymbol{A}$ 的分量系数为 0，即 $\lim\limits_{n\to\infty}f_n^{(l)}(\lambda_k)-f^{(l)}(\lambda_k)=0$ $(k=1,\ 2,\ \cdots,\ s;\ l=0,\ 1,\ \cdots,\ d_{k-1})$．　　　　证明完毕

［系 1］ 设纯量幂级数 $\sum\limits_{k=0}^{\infty}c_k\lambda^k$ 的收敛半径为 $r>0$．这时，对于全部特征值都在半径为 r 的圆内的矩阵 $\boldsymbol{A}\in C^{n\times n}(|\lambda_k|<r,\ k=1,\ 2,\ \cdots,\ s)$，矩阵幂级数

$$\sum_{k=0}^{\infty}c_k\boldsymbol{A}^k \tag{49}$$

绝对收敛．而且，当

$$f(\lambda)=\sum_{k=0}^{\infty}c_k\lambda^k \tag{50}$$

时，$f(\boldsymbol{A})$（按(1)式定义的矩阵函数）满足下式

$$f(\boldsymbol{A})=\sum_{k=0}^{\infty}c_k\boldsymbol{A}^k \tag{51}$$

（证明） 设 $f_n(\lambda)\triangleq\sum\limits_{k=0}^{n}c_k\lambda^k$，根据大家所熟悉的幂级数性质，幂级数与其微分具有相同的收敛半径，$f_n^{(l)}(\lambda)$ 对于 $|\lambda|<r$ 绝对收敛．因此，对于$|\lambda_k|<r$ 的矩阵 $\boldsymbol{A}$，由定理 3 得

$$\lim_{n\to\infty}f_n(\boldsymbol{A})=\sum_{k=0}^{\infty}c_k\boldsymbol{A}^k=f(\boldsymbol{A})$$　　　　证明完毕

［例 7］ 对于

$$(1-\lambda)^{-1}=1+\lambda+\lambda^2+\cdots+\lambda^k+\cdots\quad(r=1)$$

有

$$(\boldsymbol{I}-\boldsymbol{A})^{-1}=\boldsymbol{I}+\boldsymbol{A}+\boldsymbol{A}^2+\cdots+\boldsymbol{A}^k+\cdots$$
$$(|\lambda_k|<1,\ k=1,\ 2,\ \cdots,\ s)$$

［例 8］ 对于

$$\log(1+\lambda)=\lambda-\frac{1}{2}\lambda^2+\frac{1}{3}\lambda^3+\cdots\quad(r=1)$$

有

$$\log(\boldsymbol{I}+\boldsymbol{A})=\boldsymbol{A}-\frac{1}{2}\boldsymbol{A}^2+\frac{1}{3}\boldsymbol{A}^3+\cdots$$
$$(|\lambda_k|<1,\ k=1,\ 2,\ \cdots,\ s)$$

矩阵函数用有限级数表示(标准形 4)

因为(26)式拉格朗日-西勒维斯特内插多项式 p 最高是 $(m-1)$ 次 $(m\triangleq\deg\psi_m)$，按照 λ 的幂整理，则可以表示成

$$p(\lambda)=a_0+a_1\lambda+a_2\lambda^2+\cdots+a_{m-1}\lambda^{m-1} \tag{52}$$

因此，对于任意 $\boldsymbol{A}\in C^{n\times n}$，矩阵函数可以表示成

$$f(\boldsymbol{A})=a_0\boldsymbol{I}+a_1\boldsymbol{A}+a_2\boldsymbol{A}^2+\cdots+a_{m-1}\boldsymbol{A}^{m-1} \tag{53}$$

上式我们称之为 $f(\boldsymbol{A})$ 的标准形 4. 因 p 在 $\boldsymbol{A}$ 的谱上与 f 一致，则系数 $a_0, a_1, \cdots, a_{m-1}$ 满足下列 m 个一次方程式

$$\left.\begin{array}{l}
\left.\begin{array}{l}
f(\lambda_1)=a_0+a_1\lambda_1+\cdots+a_{m-1}\lambda_1^{m-1}\\
\vdots\\
f^{(d_1-1)}(\lambda_1)=a_{d_1-1}(d_1-1)!+a_{d_1}d_1!\lambda_1+\cdots+a_{m-1}\dfrac{(m-1)!}{(m-d_1)!}\lambda_1^{m-d_1}
\end{array}\right\}d_1\\
\vdots\\
\left.\begin{array}{l}
f(\lambda_s)=a_0+a_1\lambda_s+\cdots+a_{m-1}\lambda_s^{m-1}\\
\vdots\\
f^{(d_s-1)}(\lambda_s)=a_{d_s-1}(d_s-1)!+a_{d_s}d_s!\lambda_s+\cdots+a_{m-1}\dfrac{(m-1)!}{(m-d_s)!}\lambda_s^{m-d_s}
\end{array}\right\}d_s
\end{array}\right\}m \tag{54}$$

因为拉格朗日-西勒维斯特内插多项式只有一个存在，则 $a_0, a_1, \cdots, a_{m-1}$ 也是唯一确定的. 即(54)式的一次方程式对于 $a_0, a_1, \cdots, a_{m-1}$ 始终具有唯一解. 特别是，在 $d_1=d_2=\cdots=d_s=1$(即 $m=s$)的情况下，(54)式变成

$$\begin{bmatrix} f(\lambda_1)\\ \vdots\\ f(\lambda_s)\end{bmatrix}=\begin{bmatrix}1 & \lambda_1 & \cdots & \lambda_1^{s-1}\\ 1 & \lambda_2 & \cdots & \lambda_2^{s-1}\\ \vdots & \vdots & & \vdots\\ 1 & \lambda_s & \cdots & \lambda_s^{s-1}\end{bmatrix}\begin{bmatrix}a_0\\ a_1\\ \vdots\\ a_{m-1}\end{bmatrix} \tag{55}$$

系数矩阵行列式变成 P-I-2 中讲过的范得蒙行列式. 因此，系数矩阵是正则的.

(注意)

如关于矩阵函数的定义中所述，可以用特征多项式代替最小多项式来规定 $f(\boldsymbol{A})$. 用有限级数表示时，在这种情况下为

$$f(\boldsymbol{A})=\beta_0\boldsymbol{I}+\beta_1\boldsymbol{A}+\cdots+\beta_{n-1}\boldsymbol{A}^{n-1} \tag{56}$$

在与 $\boldsymbol{A}$ 的特征多项式有关的谱上，f 和 $(n-1)$ 次多项式

$$q(\lambda)=\beta_0+\beta_1\lambda+\cdots+\beta_{n-1}\lambda^{n-1} \tag{57}$$

一致，式(60)中的 $\beta_0, \beta_1, \cdots, \beta_{n-1}$ 可以根据这个条件确定.

[例 9]

$$\boldsymbol{A}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix},\ \psi(\lambda)=(\lambda-1)^2,\ \psi_m(\lambda)=(\lambda-1)$$

利用(56)式，因 $m=1$，则 e^A 可以表示成

$$e^A=a_0\boldsymbol{I}$$

a_0 可以由(54)式决定，$e=a_0$.

另外，利用(56)式，得

$$e^{A}=\beta_0\boldsymbol{I}+\beta_1\boldsymbol{A}$$

β_0, β_1 可以由

$$e=\beta_0+\beta_1,\ e=\beta_1$$

决定($\beta_0=0$, $\beta_1=e$)。

［例 10］ 对于

$$\boldsymbol{A}=\begin{bmatrix}0 & 0 & -2\\ 0 & 1 & 0\\ 1 & 0 & 3\end{bmatrix}$$

$$\psi(\lambda)=(\lambda-1)^2(\lambda-2),\ \psi_m(\lambda)=(\lambda-1)(\lambda-2)$$

我们利用(56)式来求 e^{At}.

$$e^{At}=\beta_0\boldsymbol{I}+\beta_1\boldsymbol{A}+\beta_2\boldsymbol{A}^2 \tag{58}$$

$$\left.\begin{aligned}e^t&=\beta_0+\beta_1+\beta_2\\ te^t&=\beta_1+2\beta_2\\ e^{2t}&=\beta_0+2\beta_1+4\beta_2\end{aligned}\right\}\Rightarrow\left.\begin{aligned}\beta_0&=-2te^t+e^{2t}\\ \beta_1&=3te^t+2e^t-2e^{2t}\\ \beta_2&=e^{2t}-e^t-te^t\end{aligned}\right\} \tag{59}$$

此外，利用(53)式，则

$$e^{At}=\alpha_0\boldsymbol{I}+\alpha_1\boldsymbol{A}$$

$$\left.\begin{aligned}e^t&=\alpha_0+\alpha_1\\ e^{2t}&=\alpha_0+2\alpha_1\end{aligned}\right\}\Rightarrow\begin{aligned}\alpha_0&=2e^t-e^{2t}\\ \alpha_1&=e^{2t}-e^t\end{aligned} \tag{60}$$

最后得

$$e^{At}=\begin{bmatrix}2e^t-e^{2t} & 0 & 2e^t-2e^{2t}\\ 0 & e^t & 0\\ e^{2t}-e^t & 0 & 2e^{2t}-e^t\end{bmatrix}$$

(注意)

在例 10 中，由求 α_0, α_1(β_0, β_1, β_2) 的过程显而易见，α_0, α_1 是线性独立函数 e^t, e^{2t} (e^t, te^t, e^{2t})的线性独立的线性组合. 因此，α_0, α_1(β_0, β_1, β_2)也是线性独立的 t 的函数. 该结论可以一般化，当表示成

$$e^{At}=\alpha_0(t)\boldsymbol{I}+\alpha_1(t)\boldsymbol{A}+\cdots+\alpha_{m-1}(t)\boldsymbol{A}^{m-1}$$

或

$$e^{At}=\beta_0(t)\boldsymbol{I}+\beta_1(t)\boldsymbol{A}+\cdots+\beta_{n-1}(t)\boldsymbol{A}^{n-1}$$

时，α_0, α_1, $\cdots$, α_{m-1}(β_0, β_1, $\cdots$, β_{n-1})是在任意时间区间上的线性独立的函数.

平方根矩阵 $A^{1/2}$

关于给定矩阵 $\boldsymbol{A}\in C^{n\times n}$ 的平方根(矩阵)的特殊情况，在 B-I-11 中已讲过. 即，当埃尔米特矩阵是准正定时，将 $\boldsymbol{A}$ 用酉阵 $\boldsymbol{P}$ 写成 $\boldsymbol{A}=\boldsymbol{P}^*\boldsymbol{\Lambda}\boldsymbol{P}$. $\boldsymbol{\Lambda}$ 是以 $\boldsymbol{A}$ 的特征值 $\lambda_i\geqslant0$ 为主对角线上的元素的实对角矩阵. 这时，$\boldsymbol{A}$ 的平方根 $\boldsymbol{A}^{1/2}$ 用

$$\boldsymbol{A}^{1/2}=\boldsymbol{P}^*\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{P} \tag{61}$$

给出. $\boldsymbol{\Lambda}^{1/2}$ 是以 λ_i 的非负平方根 $\lambda_i^{1/2}\geqslant0$ 为主对角线上元素的实对角矩阵.

下面，讲讲 $\boldsymbol{A}$ 的平方根矩阵的唯一性及与其有关的若干注意事项.

顾名思义，所谓平方根矩阵就是满足 $\boldsymbol{X}^2=\boldsymbol{A}$ 的矩阵 $\boldsymbol{X}$. 由此显而易见，除了(61)式以外，还有 $\boldsymbol{A}$ 的平方根存在(对于准正定的埃尔米特矩阵 $\boldsymbol{A}\in C^{n\times n}$). 例如，虽然 $\boldsymbol{A}^{1/2}$ 是和上述同样的矩阵，但是对于 $\boldsymbol{B}\triangleq-\boldsymbol{P}^*\boldsymbol{A}^{1/2}\boldsymbol{P}$, $\boldsymbol{B}^2=\boldsymbol{A}$ 也成立. 或者更一般化，设 $\widetilde{\boldsymbol{A}}$ 为满足 $\widetilde{\boldsymbol{A}}^2=\boldsymbol{A}$ 的任意矩阵，则 $\boldsymbol{C}\triangleq\boldsymbol{P}^*\widetilde{\boldsymbol{A}}\boldsymbol{P}$ 是 $\boldsymbol{A}$ 的平方根. 这样，平方根矩阵的唯一性就不存在了. 但是，其中和 $\boldsymbol{A}$ 一样为准正定的只有一个，这个结果在应用中很重要.

［定理 4］ 准正定埃尔米特矩阵 $\boldsymbol{A}\in C^{n\times n}$ 的准正定埃尔米特平方根矩阵唯一确定.

(证明) 显然，(61)式的 $\boldsymbol{A}^{1/2}$ 是准正定的. 设 $\boldsymbol{A}$ 的相异特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_\sigma$, 如 B-I-11 中定理 5 所示，将 $\boldsymbol{P}^*\boldsymbol{A}^{1/2}\boldsymbol{P}$ 稍作整理即可写成谱分解形

$$\boldsymbol{A}^{1/2}=\sum_{i=1}^{\sigma}\lambda_i^{1/2}\boldsymbol{E}_i \tag{62}$$

由(62)式显而易见，$\boldsymbol{A}=\sum\limits_{i=1}^{\sigma}\lambda_i\boldsymbol{E}_i$ 成立，它不外乎是 $\boldsymbol{A}$ 的谱分解形. 现在，设埃尔米特矩阵 $\boldsymbol{B}\in C^{n\times n}$ 为 $\boldsymbol{A}$ 的准正定平方根. 只要证明，实际上 $\boldsymbol{B}$ 和(62)式的 $\boldsymbol{A}^{1/2}$ 一致即可. 设 $\boldsymbol{B}$ 的谱分解为

$$\boldsymbol{B}=\sum\beta_i\boldsymbol{F}_i \tag{63}$$

因 $\boldsymbol{B}\geqslant\boldsymbol{0}$, 则 $\beta_i\geqslant0$. $\boldsymbol{B}^2$ 的谱分解形变成

$$\boldsymbol{B}^2=\sum\beta_i^2\boldsymbol{F}_i \tag{64}$$

因 $\boldsymbol{B}^2=\boldsymbol{A}$, $\boldsymbol{B}^2$ 及 $\boldsymbol{A}$ 的谱分解是唯一的(B-I-11 中定理 5)，适当更换编号后，则 $\beta_i^2=\lambda_i$, $\boldsymbol{F}_i=\boldsymbol{E}_i$ 必须成立. 因 $\beta_i\geqslant0$, $\lambda^{1/2}\geqslant0$, 则 $\beta_i=\lambda^{1/2}$, $\boldsymbol{B}$ 与(62)式的 $\boldsymbol{A}^{1/2}$ 一致.

证明完毕

(注意)

在 $\boldsymbol{A}\in C^{n\times n}$ 不一定是埃尔米特矩阵的情况下，求其平方根时要注意下面几个问题. 设 $\boldsymbol{A}$ 为正则矩阵($\det\boldsymbol{A}\neq0$). 一般，复变函数 $f(\lambda)=\lambda^{1/2}$ 是多值函数，在 $\lambda=0$ 有分枝点(branch point). 现在假定，对于 $\lambda^{1/2}$ 只考虑 $|\lambda|>0$ 及在包含 $\boldsymbol{A}$ 的所有特征值的区域上定义的分枝(branch)，所以 $\lambda^{1/2}$ 在 $\boldsymbol{A}$ 的谱上可以明确定义. 因此，在矩阵函数定义(1)的意义上，$\boldsymbol{A}^{1/2}$ 是确定的. 而且，因 $(\lambda^{1/2})^2=\lambda$, 由矩阵函数的性质(v)，则 $(\boldsymbol{A}^{1/2})^2=\boldsymbol{A}$. 此外，若 $\boldsymbol{A}$ 在 $\lambda=0$ 有 2 重特征值(考虑最小多项式包含 λ^2 因子的情况)，由于在 $\lambda=0$ 处的分枝点，则 $\lambda^{1/2}$ 在 $\boldsymbol{A}$ 的谱上不能定义. 因此，不能用定义(1)确定 $\boldsymbol{A}^{1/2}$. 与此相关，我们来看一下，是否有比用内插多项式定义的更为广泛的矩阵函数.

现在，假定 $\boldsymbol{A}$ 为单位矩阵 $\boldsymbol{I}$. 显然，对于其平方根，定义(1)只给出 $\pm\boldsymbol{I}$. 但是，满足 $\boldsymbol{X}^2=\boldsymbol{I}$ 的矩阵 $\boldsymbol{X}$ 不仅限于 $\pm\boldsymbol{I}$, 例如还包括

$$\boldsymbol{X}=\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}, \begin{bmatrix}-1 & 0\\0 & 1\end{bmatrix}.$$

而且，当 $\boldsymbol{P}\in C^{2\times2}$ 为任意正则矩阵时，

$$\boldsymbol{X}=\boldsymbol{P}^{-1}\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}\boldsymbol{P}$$

也是 $\boldsymbol{I}$ 的平方根. 由此可见，对于给定 $\boldsymbol{A}\in C^{n\times n}$, 用内插多项式(定义(1))不能求出多值函数 $\boldsymbol{A}^{1/2}$ 的全体. 因此，还可以考虑更一般的矩阵函数的定义，读者可参看文献[72].

对于上述具有 2 重特征值的矩阵 $\boldsymbol{A}$, 既使一般化定义, $\boldsymbol{A}^{1/2}$ 仍然不是确定的.

(问题 1) 设 $\boldsymbol{A}\in C^{n\times n}$, $\boldsymbol{B}\in C^{n\times n}$ 为埃尔米特矩阵, $\boldsymbol{A}\geqslant\boldsymbol{B}\geqslant\boldsymbol{0}$, 试证明 $\det\boldsymbol{A}\geqslant\det\boldsymbol{B}\geqslant 0$. 而且, 当 $\boldsymbol{A}^{1/2}$, $\boldsymbol{B}^{1/2}$ 为 $\boldsymbol{A}$, $\boldsymbol{B}$ 的埃尔米特准正定平方根时, 试证明 $\boldsymbol{A}^{1/2}\geqslant\boldsymbol{B}^{1/2}$.

在自动控制中的应用, 线性常系数系统 $\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{u}$; $\boldsymbol{y}=\boldsymbol{C}\boldsymbol{x}+\boldsymbol{D}\boldsymbol{u}$

现在, 暂且来讲讲在标题所示系统分析中的应用. 这里, $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$ 是以实数为元素的常数矩阵. 因为这种方程式是变系数系统的特殊情况, 所以在 A-I-6—A-I-8 中讲过的关于变系数系统的性质, 这里照样成立. 而且, 由于是常系数, 所以能进行更详细的分析.

[问题 1] 对应于 A-I-6 中(第 197 页)(1), (2), 应用这样的方程式情况又如何.

A-I-9　解的存在和性质

在 A-I-6 中关于变系数系统讲过的所有内容在这里都成立．根据 A-I-6 中定理 4，在这种情况下状态转移矩阵是

$$\frac{\partial \boldsymbol{\Phi}(t,\ s)}{\partial t}=\boldsymbol{A}\boldsymbol{\Phi}(t,\ s),\quad \boldsymbol{\Phi}(s,\ s)=\boldsymbol{I} \tag{1}$$

的唯一解．大家知道，$de^{\lambda t}/dt=\lambda e^{\lambda t}$，由 B-I-13 中定理 2 得

$$\frac{\partial e^{\boldsymbol{A}(t-s)}}{\partial t}=\boldsymbol{A}e^{\boldsymbol{A}(t-s)},\quad e^{\boldsymbol{A}(s-s)}=e^{0}=\boldsymbol{I} \tag{2}$$

因而可见，$e^{A(t-s)}$满足(1)式，由于解的唯一性，除此之外再没有其它解．即，在常系数系统中

$$\boldsymbol{\Phi}(t,\ s)=e^{\boldsymbol{A}(t-s)} \tag{3}$$

［问题 2］　试证明 $\det e^{At}=e^{\mathrm{tr}At}$.

利用(3)式，可以得到

$$\boldsymbol{x}(t)=\boldsymbol{s}(t;\boldsymbol{x}(t_0),\ t_0,\ \boldsymbol{u})=e^{\boldsymbol{A}(t-t_0)}\boldsymbol{x}(t_0)+\int_{t_0}^{t}e^{\boldsymbol{A}(t-\tau)}\boldsymbol{B}\boldsymbol{u}(\tau)d\tau \tag{4}$$

$$\boldsymbol{y}(t)=\boldsymbol{\rho}(t;\boldsymbol{x}(t_0),\ t_0,\ \boldsymbol{u})=\underbrace{\boldsymbol{C}e^{\boldsymbol{A}(t-t_0)}\boldsymbol{x}(t_0)}_{\text{零输入响应}}+\boldsymbol{C}\int_{t_0}^{t}\underbrace{e^{\boldsymbol{A}(t-\tau)}\boldsymbol{B}\boldsymbol{u}(\tau)d\tau+\boldsymbol{D}\boldsymbol{u}(t)}_{\text{零状态响应}} \tag{5}$$

它们相当于 A-I-6 中的(45)，(46)式．

（注意）

这里和 A-I-6 中大不相同．在常系数系统中，按照(4)，(5)式，可以用解析法计算状态和输出，其中 e^{At} 可以用后面将要讲的方法求出．而在变系数的情况下，一般不可能用解析法求出 $\boldsymbol{\Phi}(t,\ s)$.

［问题 3］　移位算子 T_τ 按

$$(T_\tau \boldsymbol{u})(t)\triangleq \boldsymbol{u}(t-\tau)$$

定义．试证明在常系数系统中

$$\boldsymbol{y}(t)=\boldsymbol{\rho}(t;\boldsymbol{x}_0,\ t_0;\boldsymbol{u})=\boldsymbol{\rho}(t+\tau;\boldsymbol{x}_0,\ t_0+\tau;T_\tau\boldsymbol{u}) \tag{6}$$

成立．

简单地讲，(6)式成立意味着“即使移动时间轴，常系数系统的输出也不会改变”这个性质称为定常(stationary)或非时不变(time-invariant)．

在定常系统中，既使将起始时间 t_0 取为 0，也不失去一般性．

下面我们来讲一些求 e^{At} 的方法．先讲解析法，后讲数值计算法．

e^{At} 的解析计算法

［其 1］　利用若当标准形法

根据 B-I-13 中(6)式，设 $\boldsymbol{A}$ 的若当标准形为 $\boldsymbol{J}=\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$，则

$$e^{At}=T^{-1}e^{Jt}T \tag{7}$$

e^{Jt} 可由 B-I-13 中[例 3](16a)—(16d)给出．因此，若 $\boldsymbol{T}$ 已知，则可由(7)式求出 e^{At}.

[问题 4]　设 $\lambda_j(1\leqslant j\leqslant\sigma)$ 为 $\boldsymbol{A}$ 的相异特征值．试证明常系数系统的零输入响应是时间函数

$$t^k e^{\lambda_j t}\quad k=0,\ 1,\ \cdots,\ d_j-1;\ \ j=1,\ \cdots,\ \sigma \tag{8}$$

的线性组合．

[其 2]　拉普拉斯变换法

因 e^{At} 是 $\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}$ 的状态转移矩阵，则为

$$\dot{\boldsymbol{X}}=\boldsymbol{A}\boldsymbol{X}\quad \boldsymbol{X}(0)=\boldsymbol{I} \tag{9}$$

的唯一解(参照 A-I-6)．对(9)式进行拉普拉斯变换，得

$$s\hat{\boldsymbol{X}}-\boldsymbol{I}=\boldsymbol{A}\hat{\boldsymbol{X}} \tag{10}$$

式中 $\hat{\boldsymbol{X}}$ 是 $\boldsymbol{X}$ 的拉普拉斯变换 $\mathscr{L}[\boldsymbol{X}]$．由(10)式，则

$$\hat{\boldsymbol{X}}=(s\boldsymbol{I}-\boldsymbol{A})^{-1}=\frac{\operatorname{adj}(s\boldsymbol{I}-\boldsymbol{A})}{\det(s\boldsymbol{I}-\boldsymbol{A})} \tag{11}$$

因此，用(11)式求出 $\hat{\boldsymbol{X}}$ 后，进行拉普拉斯反变换便可以得到 e^{At}．用(11)式求 $\hat{\boldsymbol{X}}$ 比较麻烦．此外还有更简单的方法．

如 B-I-12 定理 1 证明中所述，$\operatorname{adj}(s\boldsymbol{I}-\boldsymbol{A})$ 最多是 $(n-1)$ 次多项式，可以写成

$$\operatorname{adj}(s\boldsymbol{I}-\boldsymbol{A})=\boldsymbol{B}_0+s\boldsymbol{B}_1+\cdots+s^{n-1}\boldsymbol{B}_{n-1} \tag{12}$$

$\det(s\boldsymbol{I}-\boldsymbol{A})$ 是 s 的 n 次首一多项式，可以写成

$$\det(s\boldsymbol{I}-\boldsymbol{A})=\alpha_0+s\alpha_1+\cdots+s^{n-1}\alpha_{n-1}+s^n \tag{13}$$

现在来求 $\boldsymbol{B}_i$, α_i．由 B-I-11 中(11)式得

$$(s\boldsymbol{I}-\boldsymbol{A})(\boldsymbol{B}_0+s\boldsymbol{B}_1+\cdots+s^{n-1}\boldsymbol{B}_{n-1})=(\alpha_0+s\alpha_1+\cdots+s^{n-1}\alpha_{n-1}+s^n)\boldsymbol{I}. \tag{14}$$

[定理 1]　α_i, $\boldsymbol{B}_i$ 可以用下列迭代公式求出．

$$\begin{aligned}
&\boldsymbol{B}_{n-1}=\boldsymbol{I}, && \alpha_{n-1}=-\operatorname{tr}\boldsymbol{A}\\
&\boldsymbol{B}_{n-2}=\boldsymbol{A}\boldsymbol{B}_{n-1}+\alpha_{n-1}\boldsymbol{I}, && \alpha_{n-2}=-\frac{1}{2}\operatorname{tr}(\boldsymbol{B}_{n-2}\boldsymbol{A})\\
&\boldsymbol{B}_{n-3}=\boldsymbol{A}\boldsymbol{B}_{n-2}+\alpha_{n-2}\boldsymbol{I}, && \alpha_{n-3}=-\frac{1}{3}\operatorname{tr}(\boldsymbol{B}_{n-3}\boldsymbol{A})\\
&\quad\vdots && \quad\vdots\\
&\boldsymbol{B}_{n-k}=\boldsymbol{A}\boldsymbol{B}_{n-k+1}+\alpha_{n-k+1}\boldsymbol{I}, && \alpha_{n-k}=-\frac{1}{k}\operatorname{tr}(\boldsymbol{B}_{n-k}\boldsymbol{A})\\
&\quad\vdots && \quad\vdots\\
&\boldsymbol{B}_1=\boldsymbol{A}\boldsymbol{B}_2+\alpha_2\boldsymbol{I}, && \alpha_1=-\frac{1}{n-1}\operatorname{tr}(\boldsymbol{B}_1\boldsymbol{A})\\
&\boldsymbol{B}_0=\boldsymbol{A}\boldsymbol{B}_1+\alpha_1\boldsymbol{I}, && \alpha_0=-\frac{1}{n}\operatorname{tr}(\boldsymbol{B}_0\boldsymbol{A})\\
&(\boldsymbol{0}=\boldsymbol{A}\boldsymbol{B}_0+\alpha_0\boldsymbol{I}).
\end{aligned} \tag{15}$$

将(14)式两边 s 的同次项彼此比较，便可得到上列各个 $\boldsymbol{B}_{n-k}$ 的计算式．α_{n-k} 式的证明示于本章附录．不用上面括号内的算式也可以求出 $\boldsymbol{B}_i$, α_i, 该式用于验算所求出的结果．

上述求 $\boldsymbol{B}_{n-k}$, α_{n-k} 的方法称为 Leverrier-Faddeeva 算法，有时也称为 Souriau-Frame

方法[24,31].

对于 $\hat{X}$ 的表示，部分分式展开式也是有用的．根据 B-I-13 中(41)，(42)式，

$$[\lambda \boldsymbol{I}-\boldsymbol{A}]^{-1}=\sum_{k=1}^{\sigma}\left[\frac{\boldsymbol{R}_{k0}}{s-\lambda_k}+\frac{\boldsymbol{R}_{k1}}{(s-\lambda_k)^2}+\cdots+\frac{\boldsymbol{R}_{kd_k-1}}{(s-\lambda_k)^{d_k}}\right] \tag{16}$$

式中 $\boldsymbol{R}_{kl}$ 是常数矩阵，可以用

$$\boldsymbol{R}_{kl}=\frac{1}{(d_k-l)!}\left[\frac{d^{(d_k-1)}}{ds^{(d_k-1)}}\left(\frac{(s-\lambda_k)^{d_k}\boldsymbol{R}(s)}{\psi_m(s)}\right)\right]_{s=\lambda_k} \tag{17}$$

给出．将(16)式进行拉普拉斯反变换，容易得

$$e^{At}=\sum_{k=1}^{\sigma}\left[\boldsymbol{R}_{k0}+t\boldsymbol{R}_{k1}+\cdots+\frac{t^{d_k-1}}{(d_k-1)!}\boldsymbol{R}_{kd_k-1}\right]e^{\lambda_k t} \tag{18}$$

[其 3]　有限级数表示

由 B-I-13 中(52)—(54)式可以得到下列求 e^{At} 的方法．设 $\boldsymbol{A}$ 的最小多项式的次数为 m，考虑 $m-1$ 次多项式

$$\gamma(\lambda)\triangleq a_0(t)+a_1(t)\lambda+\cdots+a_{m-1}(t)\lambda^{m-1} \tag{19}$$

根据满足

$$\left.\frac{d^k}{d\lambda^k}e^{\lambda t}\right|_{\lambda=\lambda_j}=\left.\frac{d^k}{d\lambda^k}\gamma(\lambda)\right|_{\lambda=\lambda_j}\quad k=0,1,\cdots,d_j-1;\ j=1,\cdots,\sigma \tag{20}$$

的条件确定 $a_i(t)$ $(i=0,1,\cdots,m-1)$．于是，e^{At} 可由

$$e^{At}=\gamma(\boldsymbol{A})=a_0(t)\boldsymbol{I}+a_1(t)\boldsymbol{A}+\cdots+a_{m-1}(t)\boldsymbol{A}^{m-1} \tag{21}$$

给出．

e^{At} 的数值计算法

在具体确定状态方程式 $\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{u}$ 的解时，就产生了计算 e^{At} 的必要性．如已讲过，该解一般可由

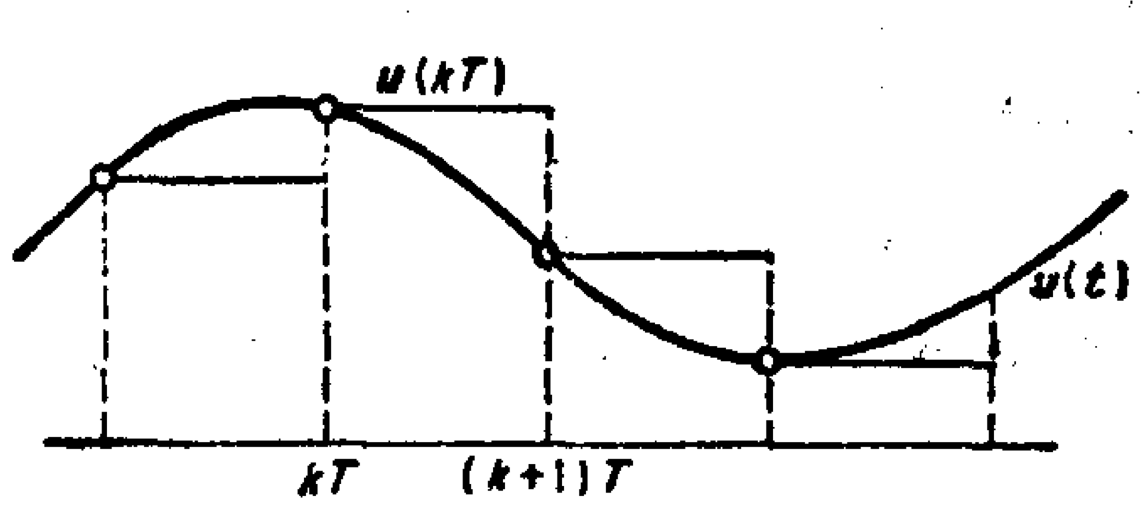
$$\boldsymbol{x}(t)=e^{A(t-t_0)}\boldsymbol{x}(t_0)+\int_{t_0}^{t}e^{A(t-\tau)}\boldsymbol{B}\boldsymbol{u}(\tau)d\tau \tag{4}$$

给出．现在我们来讨论求 $\boldsymbol{x}(t)$ 在采样时刻 $t=kT$ $(k=0,1,2,\cdots)$ 时数值的数值计算法．在(4)式中，令 $t=(k+1)T$，$t_0=kT$，则

$$\boldsymbol{x}[(k+1)T]=e^{AT}\boldsymbol{x}(kT)+e^{A(k+1)T}\int_{kT}^{(k+1)T}e^{-A\tau}\boldsymbol{B}\boldsymbol{u}(\tau)d\tau \tag{22}$$

如图 1 所示，现假定在采样区间 $[kT,(k+1)T]$ 上输入值一定(阶梯形逼近)，因

$$\boldsymbol{u}(t)=\boldsymbol{u}(kT),\quad kT\leqslant t<(k+1)T$$



$\boldsymbol{u}(t)\simeq\boldsymbol{u}(kT)\quad kT\leq t<(k+1)T$

图 1　输入用阶梯形曲线逼近

则(22)式的第二项与 $\boldsymbol{u}(kT)$ 成线性关系，得到

$$\boldsymbol{x}[(k+1)T]=e^{At}\boldsymbol{x}(kT)+\boldsymbol{G}\boldsymbol{u}(kT),$$
$$\left(\boldsymbol{G}\triangleq\int_0^T e^{A\tau}\boldsymbol{B}d\tau\right) \tag{23}$$

当 $\boldsymbol{x}(0)$ 给出时，按该式可以逐次计算出 $\boldsymbol{x}(kT)$, $k=1, 2, \cdots$. 采样时刻的 $\boldsymbol{x}(kT)$ 求出后，输出方程式 $\boldsymbol{y}=\boldsymbol{C}\boldsymbol{x}+\boldsymbol{D}\boldsymbol{u}$ 的解也可以由

$$\boldsymbol{y}(kT)=\boldsymbol{C}\boldsymbol{x}(kT)+\boldsymbol{D}\boldsymbol{u}(kT) \tag{24}$$

计算出.

在(23)式的右边，包含 e^{At} 和 $\boldsymbol{G}$. 以后将要讲到，$\boldsymbol{G}$ 与 e^{At} 的计算基本上是一样的，所以现在只讲 e^{At} 的数值计算.

e^{At} 的数值解和解析计算法不同，不能假定 $\boldsymbol{A}$ 的特征值及其若当标准形是已知的. 现在知道的计算方法有：(i) 取 e^{AT} 的幂级数表示 $\sum_{k=1}^{\infty}(\boldsymbol{A}T)^k/k!$ 的有限项的方法[73]. (ii)巧妙地利用凯莱-哈密顿 (Cayley-Hamilton) 定理 (B-I-12 定理 1) 及 Leverrier-Faddeeva 算法(参照(15)式)等矩阵固有性质的方法[74-76]. (iii)利用拉普拉斯变换 $\mathscr{L}[e^{At}]=[s\boldsymbol{I}-\boldsymbol{A}]^{-1}$(参照其 2, e^{At} 的解析计算法)的方法等. 其中，方法(i)在原理上最简单. 设 e^{At} 幂级数表示的前 N 项和为 $\boldsymbol{M}$, 余下部分为 $\boldsymbol{R}$, 当 N 取得足够大时, $e^{At}\approx\boldsymbol{M}$ 可以达到任意精度.

问题是，为了满足给定的误差范围，N 确定到多大才可以，必须注意给出的 N 不要过大[78,79]. 一般，因为在这个方法中必须计算 $\boldsymbol{A}$ 的幂 $\boldsymbol{A}^k(k=0, 1, 2, \cdots, N-1)$, 所以计算量比方法(ii)大. 在拉普拉斯变换法(iii)中，不是从 $[s\boldsymbol{I}-\boldsymbol{A}]^{-1}$ 的拉普拉斯反变换中求 e^{At}, 而是将 e^{AT} 展成周期为 $2T$ 的傅里叶级数，其傅里叶系数由关系式 $[s\boldsymbol{I}-\boldsymbol{A}]\mathscr{L}[e^{At}]=\boldsymbol{I}$ 决定. 当然，实际上傅里叶级数也是用有限项逼近. 总的说来，该方法的计算量也和方法(i)差不多.

我们主要对利用矩阵特有性质这一点感兴趣，下面来介绍方法(ii)，特别是其中利用凯莱-哈密顿定理的方法[74,75,79]. 方法(ii)的计算量也比前面讲的两种小.

该方法原理上也和方法(i)一样，将 e^{At} 表示成幂级数进行计算. 但不同的是，根据凯莱-哈密顿定理，对于 $n\times n$ 矩阵 $\boldsymbol{A}$, 其幂 $\boldsymbol{A}^k(k\geqslant n)$ 可以用 $\boldsymbol{I}, \boldsymbol{A}, \cdots, \boldsymbol{A}^{n-1}$ 的线性组合表示，利用该性质，$\boldsymbol{A}^n$ 以上的幂都不需直接计算. 为了利用凯莱-哈密顿定理，必须先求出 $\boldsymbol{A}$ 的特征多项式，但后者也可以利用 Leverrier 公式简单地由 $\boldsymbol{I}, \boldsymbol{A}, \cdots, \boldsymbol{A}^n$ 决定[24]. Mastascusa 提出将凯莱-哈密顿定理和该公式结合起来的方案，下面讲的就是这个方法.

设 $\boldsymbol{A}$ 的特征多项式为

$$\psi(\lambda)\triangleq\det[\lambda\boldsymbol{I}-\boldsymbol{A}]\triangleq\lambda^n+\alpha_{n-1}\lambda^{n-1}+\cdots+\alpha_0 \tag{25}$$

由凯莱-哈密顿定理，$\psi(\boldsymbol{A})=\boldsymbol{0}$, 改写后得

$$\boldsymbol{A}^n=-\alpha_{n-1}\boldsymbol{A}^{n-1}-\alpha_{n-2}\boldsymbol{A}^{n-2}-\cdots-\alpha_0\boldsymbol{I} \tag{26}$$

反复使用(26)式，可见 $\boldsymbol{A}$ 的幂 $\boldsymbol{A}^{n+m}(m\geqslant 0)$ 一般也可以由 $\boldsymbol{I}, \boldsymbol{A}, \cdots, \boldsymbol{A}^{n-1}$ 线性表出，即可写成

$$\boldsymbol{A}^{n+m}=\alpha_{0m}\boldsymbol{I}+\alpha_{1m}\boldsymbol{A}+\cdots+\alpha_{n-1,m}\boldsymbol{A}^{n-1}(m=0, 1, 2, \cdots) \tag{27}$$

右边系数可按如下方法决定. 假定 $\boldsymbol{A}^{n+m-1}$ 的系数 $\alpha_{0m-1}, \alpha_{1m-1}, \cdots, \alpha_{n-1,m-1}$ 为已知，因

$A^{n+m}=A^{n+m-1}A$, 将(26), (27)式代入该式, 两边进行比较则得

$$\begin{aligned}\alpha_{0m}&=-\alpha_0\alpha_{n-1,m-1}\\ \alpha_{1m}&=\alpha_{0m-1}-\alpha_1\alpha_{n-1}\alpha_{n-1,m-1}\\ &\vdots\\ \alpha_{n-1,m}&=\alpha_{n-2,m-1}-\alpha_{n-1}\alpha_{n-1,m-1}\end{aligned} \tag{28}$$

因 $\alpha_{00}=-\alpha_0$, $\alpha_{10}=-\alpha_1$, $\cdots$, $\alpha_{n-1,0}=-\alpha_{n-1}$, 则从 $m=1$ 起依次使用(28)式可以决定 α_{0m}, α_{1m}, $\cdots$, $\alpha_{n-1,m}$. 那么, 将(27)式代入 e^{AT} 的幂级数表示式, 得

$$\begin{aligned}e^{AT}&=\sum_{k=0}^{\infty}\frac{1}{k!}(AT)^k=\sum_{k=0}^{n-1}\frac{T^k}{k!}A^k+\sum_{m=0}^{\infty}\frac{T^{n+m}}{(n+m)!}\sum_{k=0}^{n-1}\alpha_{km}A^k\\ &=\sum_{k=0}^{n-1}A^k\left[\frac{T^k}{k!}+\sum_{m=0}^{\infty}\frac{\alpha_{km}T^{n+m}}{(n+m)!}\right]\end{aligned} \tag{29}$$

在 e^{AT} 的幂级数表示式中取到第 $(N+1)$ 项, 相当于在(29)式[ ]内的第二项取到 $m=N-n$. 使用(29)式计算 e^{AT} 的步骤如下:

(i) 计算 A^2, A^3, $\cdots$, A^n.

(ii) 按下列的 Leverrier 公式(也称为 Bôcher 公式, 其证明见本章附录)确定 A 的特征多项式 $\psi(\lambda)$.

$$\begin{aligned}\alpha_{n-1}&=-\operatorname{tr}(A)\\ \alpha_{n-2}&=-\frac{1}{2}[\alpha_{n-1}\operatorname{tr}(A)+\operatorname{tr}(A^2)]\\ &\vdots\\ \alpha_0&=-\frac{1}{n}[\alpha_1\operatorname{tr}(A)+\alpha_2\operatorname{tr}(A^2)+\cdots+\alpha_{n-1}\operatorname{tr}(A^{n-1})+\operatorname{tr}(A^n)]\end{aligned} \tag{30}$$

(iii) 根据给定的精度, 在(29)式[ ]内第二项中适当地选取有限项, 在此基础上计算 e^{AT}.

为了按近似式(23)解出状态方程式, 除了 e^{AT} 之外还必需计算 $G=\int_0^T e^{A\tau}Bd\tau$. 但因

$$\int_0^T e^{A\tau}d\tau=\sum_{k=0}^{\infty}\int_0^T\frac{1}{k!}A^k\tau^k d\tau=T\sum_{k=0}^{\infty}\frac{1}{(k+1)!}(AT)^k$$

则 G 的计算也可以和 e^{AT} 同样进行, 其说明这里省略.

附录

Leverrier (Bocher)公式及 Leverrier-Faddeeva 算法的证明

设 A 的特征多项式为 $\psi(\lambda)=\det[\lambda I-A]=\lambda^n+\alpha_{n-1}\lambda^{n-1}+\alpha_{n-2}\lambda^{n-2}+\cdots+\alpha_0$, 其根, 即 A 的特征值(包含重复的)为 λ_1, λ_2, $\cdots$, λ_n.

现在. 若令

$$\lambda_1^k+\lambda_2^k+\cdots+\lambda_n^k\triangleq s_k \tag{A.1}$$

由根与系数关系的牛顿公式, 则

$$-k\alpha_{n-k}=s_k+\alpha_{n-1}s_{k-1}+\cdots+\alpha_{n-k+1}s_1 \tag{A.2}$$

成立.

另外, 设 A 的若当标准形为 J, 若令 $A=TJT^{-1}$, 则 $A^k=TJ^kT^{-1}$. 而且, 根据 P-I-2 中迹的性质(iii), 一般, 因 $\operatorname{tr}(AB)=\operatorname{tr}(BA)$, 则

$$\mathrm{tr}\boldsymbol{A}^k=\mathrm{tr}[\boldsymbol{T}(\boldsymbol{J}^k\boldsymbol{T}^{-1})]=\mathrm{tr}[(\boldsymbol{J}^k)(\boldsymbol{T}\boldsymbol{T}^{-1})]=\mathrm{tr}\boldsymbol{J}^k=s_k \quad (k=1,\ 2,\ \cdots,\ n) \tag{A.3}$$

成立．因此，由(A.1)，(A.2)，(A.3)式得

$$\alpha_{n-k}=-\frac{1}{k}[\mathrm{tr}(\boldsymbol{A}^k)+\alpha_{n-1}\mathrm{tr}(\boldsymbol{A}^{k-1})+\cdots+\alpha_{n-k+1}\mathrm{tr}(\boldsymbol{A})] \quad k=1,\ 2,\ \cdots,\ n \tag{A.4}$$

于是，Leverrier 公式得证．

此外，在 e^{At} 的解析计算法[其 2]中讲的 $\boldsymbol{B}_0$, $\boldsymbol{B}_1$, $\cdots$, $\boldsymbol{B}_{n-1}$ 和(A.4)式有密切关系．现设

$$[\lambda\boldsymbol{I}-\boldsymbol{A}]^{-1}=\frac{\mathrm{adj}[\lambda\boldsymbol{I}-\boldsymbol{A}]}{\psi(\lambda)}$$

如已讲过，则可表示成

$$\mathrm{adj}[\lambda\boldsymbol{I}-\boldsymbol{A}]=\lambda^{n-1}\boldsymbol{B}_{n-1}+\lambda^{n-2}\boldsymbol{B}_{n-2}+\cdots+\boldsymbol{B}_0.$$

若特征多项式 $\psi(\lambda)$ 的系数已知，则系数矩阵 $\boldsymbol{B}_k(k=0,\ 1,\ \cdots,\ n-1)$ 可以根据(15)式左栏关系

$$\begin{aligned}
\boldsymbol{B}_{n-1}&=\boldsymbol{I}\\
\boldsymbol{B}_{n-2}&=\boldsymbol{A}\boldsymbol{B}_{n-1}+\alpha_{n-1}\boldsymbol{I}\\
&\ \vdots\\
\boldsymbol{B}_1&=\boldsymbol{A}\boldsymbol{B}_2+\alpha_2\boldsymbol{I}\\
\boldsymbol{B}_0&=\boldsymbol{A}\boldsymbol{B}_1+\alpha_1\boldsymbol{I}
\end{aligned} \tag{A.5}$$

计算．注意，由(A.5)式，则

$$\begin{aligned}
\boldsymbol{B}_{n-k-1}&=\boldsymbol{A}\boldsymbol{B}_{n-k}+\alpha_{n-k}\boldsymbol{I}=\boldsymbol{A}(\boldsymbol{A}\boldsymbol{B}_{n-k+1}+\alpha_{n-k+1}\boldsymbol{I})+\alpha_{n-k}\boldsymbol{I}=\cdots\cdots\\
&=\boldsymbol{A}^k+\alpha_{n-1}\boldsymbol{A}^{k-1}+\cdots+\alpha_{n-k+1}\boldsymbol{A}+\alpha_{n-k}\boldsymbol{I}
\end{aligned} \tag{A.6}$$

成立．

由(A.4)，(A.6)可见，特征多项式系数利用矩阵 $\boldsymbol{B}_k$ 也可以表示成

$$\alpha_{n-k}=-\frac{1}{k}\mathrm{tr}(\boldsymbol{A}\boldsymbol{B}_{n-k}) \quad (k=1,\ 2,\ \cdots,\ n) \tag{A.7}$$

因 $\boldsymbol{A}$ 和 $\boldsymbol{B}_k$ $(k=0,\ 1,\ \cdots,\ n-1)$ 可换，则(A.7)式与(15)式右栏结果相同．

A-I-10 输入输出关系

在这里，A-I-7 的结果全部成立. 根据 A-I-9 中(5)式，常系数系统脉冲响应矩阵

$$\boldsymbol{H}(t,\ \tau)=\begin{cases}\boldsymbol{C}e^{\boldsymbol{A}(t-\tau)}\boldsymbol{B}+\boldsymbol{D}\delta(t-\tau) & t\geqslant\tau\\ \boldsymbol{0} & t<\tau\end{cases}\tag{1}$$

由此显而易见，$\boldsymbol{H}(t,\ \tau)$仅是$(t-\tau)$的函数，可以写成$\boldsymbol{H}(t-\tau)$. 因此，在常系数系统的情况下，脉冲响应矩阵可以表示成

$$\boldsymbol{H}(t)=\begin{cases}\boldsymbol{C}e^{\boldsymbol{A}t}\boldsymbol{B}+\boldsymbol{D}\delta(t) & t\geqslant 0\\ \boldsymbol{0} & t<0\end{cases}\tag{2}$$

以下经常使用这个表示式.

将(2)式两边进行拉氏变换，利用 A-I-9 中(11)式，得

$$\hat{\boldsymbol{H}}(s)=\boldsymbol{C}[s\boldsymbol{I}-\boldsymbol{A}]^{-1}\boldsymbol{B}+\boldsymbol{D}=\frac{\boldsymbol{C}\operatorname{adj}[s\boldsymbol{I}-\boldsymbol{A}]\boldsymbol{B}}{\det(s\boldsymbol{I}-\boldsymbol{A})}+\boldsymbol{D}\tag{3}$$

式中 $\hat{\boldsymbol{H}}(s)\triangleq\mathscr{L}[\boldsymbol{H}(t)]$.

由 A-I-7 中(3)式和上述(1)，(2)式可见，系统的零状态响应

$$\boldsymbol{y}_{zs}(t)=\int_0^t\boldsymbol{H}(t-\tau)\boldsymbol{u}(\tau)d\tau\quad t\geqslant 0\tag{4}$$

注意，该式右边是 $\boldsymbol{H}$ 和 $\boldsymbol{u}$ 的褶积，则

$$\hat{\boldsymbol{y}}_{zs}(s)=\hat{\boldsymbol{H}}(s)\hat{\boldsymbol{u}}(s)\tag{5}$$

式中 $\hat{\boldsymbol{y}}_{zs}$，$\hat{\boldsymbol{u}}$ 分别是 $\boldsymbol{y}_{zs}(t)$，$\boldsymbol{u}(t)$的拉氏变换. 因(5)式成立，则仿照单输入单输出系统的情况，称 $\hat{\boldsymbol{H}}(s)$为传递函数矩阵. 传递函数矩阵是体 $Q(s,\ R)$上(即以实系数的 s 的有理函数为元素)的 $m\times r$ 矩阵. 设 $\boldsymbol{A}$ 的最小多项式为 $\psi_m(s)$，$\operatorname{adj}(s\boldsymbol{I}-\boldsymbol{A})$ 的全元素最大公约多项式为 $d(s)$，则由 B-I-12 中系 3 得

$$\det(s\boldsymbol{I}-\boldsymbol{A})=\psi_m(s)d(s)\tag{6}$$

即 $\det(s\boldsymbol{I}-\boldsymbol{A})$可用 $d(s)$整除. 因此，可以按下列约去分子分母的共同项

$$[s\boldsymbol{I}-\boldsymbol{A}]^{-1}=\frac{\operatorname{adj}(s\boldsymbol{I}-\boldsymbol{A})}{\det(s\boldsymbol{I}-\boldsymbol{A})}=\frac{\boldsymbol{R}(s)d(s)}{\psi_m(s)d(s)}=\frac{\boldsymbol{R}(s)}{\psi_m(s)}.\tag{7}$$

根据 $d(s)$的定义，$\boldsymbol{R}(s)$和 $\psi_m(s)$不能再约简，由(3)，(7)式得

$$\hat{\boldsymbol{H}}(s)=\frac{\boldsymbol{C}\boldsymbol{R}(s)\boldsymbol{B}}{\psi_m(s)}+\boldsymbol{D}\tag{8}$$

在这里，$\boldsymbol{C}\boldsymbol{R}(s)\boldsymbol{B}$ 和 $\psi_m(s)$有可能再约简，这与系统的可控性、可观测性密切相关，以后有机会再详细叙述.

由(3)(或(8))式得

$$\lim_{s\to\infty}\hat{\boldsymbol{H}}(s)=\boldsymbol{D}=\text{与 } s \text{ 无关的常数矩阵}\tag{9}$$

(为何?). 一般，当 $Q(s,\ R)$上的矩阵满足关系式(9)时，称该矩阵是适宜(proper)的. 当满足

$$\lim_{s\to\infty}\hat{\boldsymbol{H}}(s)=\boldsymbol{0}=\text{零矩阵} \tag{10}$$

时，称为是真正适宜(strictly proper)的.

状态空间中的坐标(基底)变换

和 A-1-7 中一样，在常系数系统中也可以按 $\tilde{\boldsymbol{x}}(t)=\boldsymbol{T}(t)\boldsymbol{x}(t)$ 引入新的变量. 我们特别感兴趣的是变换后的系统也是常系数，即 $\boldsymbol{T}$ 是与 t 无关的常数矩阵的情况. 现在，按

$$\tilde{\boldsymbol{x}}(t)=\boldsymbol{T}\boldsymbol{x}(t)\quad \boldsymbol{T}\in R^{n\times n}\quad \det\boldsymbol{T}\neq 0 \tag{11}$$

定义 $\tilde{\boldsymbol{x}}$. 当对于 $\boldsymbol{x}$

$$\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{u},\quad \boldsymbol{y}=\boldsymbol{C}\boldsymbol{x}+\boldsymbol{D}\boldsymbol{u} \tag{12}$$

成立时，对于新的变量 $\tilde{\boldsymbol{x}}$

$$\dot{\tilde{\boldsymbol{x}}}=\tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}}+\tilde{\boldsymbol{B}}\boldsymbol{u},\quad \boldsymbol{y}=\tilde{\boldsymbol{C}}\tilde{\boldsymbol{x}}+\tilde{\boldsymbol{D}}\boldsymbol{u} \tag{13}$$

$$\tilde{\boldsymbol{A}}=\boldsymbol{T}\boldsymbol{A}\boldsymbol{T}^{-1},\ \tilde{\boldsymbol{B}}=\boldsymbol{T}\boldsymbol{B},\ \tilde{\boldsymbol{C}}=\boldsymbol{C}\boldsymbol{T}^{-1},\ \tilde{\boldsymbol{D}}=\boldsymbol{D} \tag{14}$$

成立.

[问题 1] 设(13)式系统的脉冲响应矩阵为 $\tilde{\boldsymbol{H}}(t)$，试证明

$$\tilde{\boldsymbol{H}}(t)=\boldsymbol{H}(t) \tag{15}$$

现在，设 $\boldsymbol{T}$ 的第 i 列为 $\boldsymbol{t}_i$，即 $\boldsymbol{T}\triangleq[\boldsymbol{t}_1,\ \cdots,\ \boldsymbol{t}_n]$. 因 $\boldsymbol{T}$ 是正则的，故 $\boldsymbol{t}_i(i=1,\ \cdots,\ n)$ 是线性独的，是空间 R^n 的基底[1].

(11) 式可以写成

$$\boldsymbol{e}_1\tilde{x}_1+\cdots+\boldsymbol{e}_n\tilde{x}_n=\boldsymbol{t}_1x_1+\cdots+\boldsymbol{t}_nx_n \tag{11'}$$

该式表示，当系统的状态用 R^n 的元表示时，使用基底 $\boldsymbol{e}_1,\ \cdots,\ \boldsymbol{e}_n$[2]，则分量是 $\tilde{x}_1,\ \cdots,\ \tilde{x}_n$，使用基底 $\boldsymbol{t}_1,\ \cdots,\ \boldsymbol{t}_n$，则分量是 $x_1,\ \cdots,\ x_n$，即(11)式的变换是 R^n 中基底的变换. 换句话说，是坐标变换. 这一点，对于 $n=2$ 的情况，由图 2 一目了然.

下面我们来讨论，对于给定的 $\boldsymbol{A}$，选择适当的 $\boldsymbol{T}$，将得到简单的 $\tilde{\boldsymbol{A}}$.

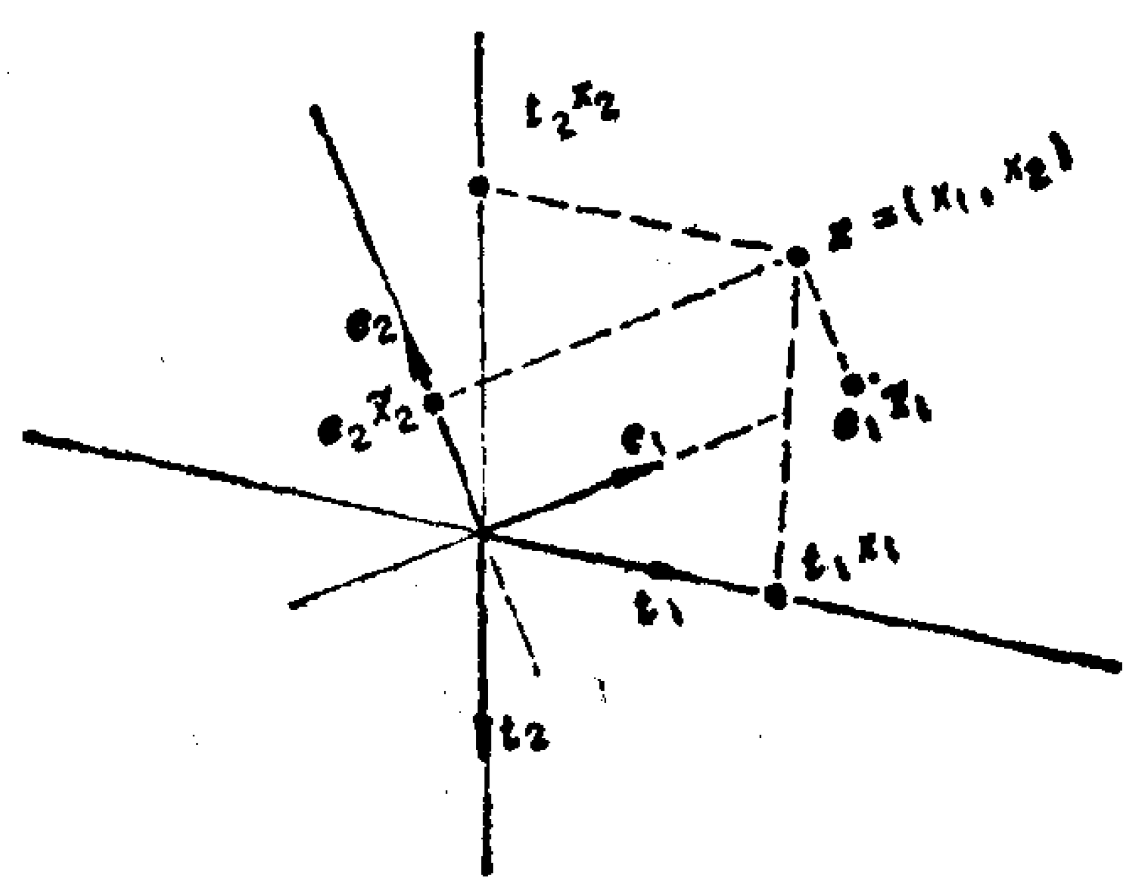


图 2

1) 在向量空间 V 中，当某一组向量 $\boldsymbol{x}_i(i=1,\ \cdots,\ n)\in V$ 是线性独立的，且任意 $\boldsymbol{x}\in V$ 都可以用 $\boldsymbol{x}_i$ 线性表出时，$\boldsymbol{x}_i$ 便称为 V 的基底(第 199 页注解). 现在 $\boldsymbol{t}_i(i=1,\ \cdots,\ n)$ 是线性独立的，对于任意 $\boldsymbol{y}\in R^n$，$\boldsymbol{T}\boldsymbol{x}=\boldsymbol{y}$ 的解必然存在，而且是唯一确定的(因 $\boldsymbol{T}$ 是正则的). 即有使 $\sum_{i=1}^{n}\boldsymbol{t}_ix_i=\boldsymbol{y}$ 成立的 x_i 存在，所以 $\boldsymbol{t}_i(i=1,\ \cdots,\ n)$ 是 R^n 的基底. 这里，x_i 称为关于 $\boldsymbol{y}$ 的基底 $\boldsymbol{t}_i$ 的分量.

2) 它称为自然基底(natural basis)或基本基底.

[其 1]　$\boldsymbol{A}$ 的特征值都是实数，且 $\boldsymbol{A}$ 为单纯矩阵的情况.

因 $\boldsymbol{A}$ 为单纯矩阵，则有 n 个线性独立的左特征向量 $\boldsymbol{y}_i^T(i=1, \cdots, n)$ 存在 (参看 B-I-11 中系 15, 定理 4). 因 $\boldsymbol{A}$ 的特征值均为实数，则 $\boldsymbol{y}_i$ 也可以认为是实向量 ($\boldsymbol{y}_i \in R^n$). 现对 $\boldsymbol{T}$ 采用

$$\boldsymbol{T}=\begin{bmatrix} \boldsymbol{y}_1^T \\ \vdots \\ \boldsymbol{y}_n^T \end{bmatrix} \in R^{n\times n} \tag{16}$$

则 $\tilde{\boldsymbol{A}}=\boldsymbol{TAT}^{-1}=\boldsymbol{\Lambda}$ (以 $\boldsymbol{A}$ 的特征值 $\lambda_1, \cdots, \lambda_n$ 为主对角线上元素的实对角矩阵) 成立 (试证明). 将其代入 (13) 式的第 1 式，得

$$\dot{\tilde{\boldsymbol{x}}}=\boldsymbol{\Lambda}\tilde{\boldsymbol{x}}+\tilde{\boldsymbol{B}}\boldsymbol{u} \tag{17}$$

即若令 $\tilde{\boldsymbol{B}}$ 的第 i 行为 $\tilde{\boldsymbol{b}}_i$，则

$$\dot{\tilde{x}}_i=\lambda\tilde{x}_i+\tilde{\boldsymbol{b}}_i\boldsymbol{u} \quad i=1, \cdots, n \tag{18}$$

因为上式是纯量微分方程式，则立即可以解出，得

$$\tilde{x}_i(t)=e^{\lambda_i t}x_i(0)+\int_0^t e^{\lambda_i(t-\tau)}\tilde{\boldsymbol{b}}_i\boldsymbol{u}(\tau)d\tau \quad i=1, \cdots, n \tag{19}$$

由此可见，第 i 个分量 $\tilde{x}_i$ 和其它分量 $\tilde{x}_j(j\neq i)$ 是相互独立的，仅决定于起始值 $\tilde{x}_i(0)$ 和输入 $\boldsymbol{u}$. 该 $\tilde{x}_i$ 称为系统的第 i 个模 (mode).

[其 2]　$\boldsymbol{A}$ 的特征值均为实数，$\boldsymbol{A}$ 不是单纯矩阵的情况.

因 $\boldsymbol{A}$ 不是单纯矩阵，则 $\tilde{\boldsymbol{A}}$ 不能化成[其 1]那样的对角矩阵. 但是适当地选择 $\boldsymbol{T}$, 可将 $\tilde{\boldsymbol{A}}$ 化成若当标准形，而且对于这种 $\boldsymbol{T}$ 可采用实矩阵，假定 $\tilde{\boldsymbol{A}}=\boldsymbol{TAT}^{-1}=\boldsymbol{J}$ (若当形). 我们来考虑一下最简单的

$$\boldsymbol{J}=\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} \lambda\in R \tag{20}$$

的情况. 根据 B-I-13(16d) 式，因

$$e^{Jt}=\begin{bmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{bmatrix}$$

则

$$\begin{aligned} \tilde{x}_1(t)&=e^{\lambda t}\tilde{x}_1(0)+te^{\lambda t}\tilde{x}_2(0)+\int_0^t e^{\lambda(t-\tau)}\tilde{\boldsymbol{b}}_1\boldsymbol{u}(\tau)d\tau+\int_0^t (t-\tau)e^{\lambda(t-\tau)}\tilde{\boldsymbol{b}}_2\boldsymbol{u}(\tau)d\tau \\ \tilde{x}_2(t)&=e^{\lambda t}\tilde{x}_2(0)+\int_0^t e^{\lambda(t-\tau)}\tilde{\boldsymbol{b}}_2\boldsymbol{u}(\tau)d\tau \end{aligned} \tag{21}$$

在这里可以看到，特别是若假定 $\boldsymbol{u}=\boldsymbol{0}$, 则无论怎样选择起始条件，都不能只激励(影响) $\tilde{x}_2$. 即，这种情况下 $\tilde{x}_1$ 和 $\tilde{x}_2$ 不是完全独立的. 对于更一般的若当形也同样可以证明，属于同一子块 $\boldsymbol{J}_{jk}$ 的分量不是独立的.

[其 3]　$\boldsymbol{A}$ 的特征值包含复数的情况

既使 $\boldsymbol{A}$ 的特征值包含有复数，也可以将 $\boldsymbol{A}$ 化成对角形或若当形，只是在这种情况下 $\boldsymbol{T}$, $\tilde{\boldsymbol{A}}$ 不是实矩阵，一般 $\boldsymbol{T}$, $\tilde{\boldsymbol{A}}\in C^{n\times n}$. 因为实际系统中的变量都是实数，所以在方程式中出现复数不大方便. 因此，我们要研究一下，在实数范围内 $\tilde{\boldsymbol{A}}$ 会变成什么形状. 首先必须注意，因 $\boldsymbol{A}$ 是实矩阵，则 $\det(\lambda\boldsymbol{I}-\boldsymbol{A})=0$ 是实系数代数方程式. 因此，若复数 $\lambda(\mathrm{Im}\lambda\neq0)$ 是该方程式的根，则 $\bar{\lambda}$ (λ 的共轭复数) 也是它的根. 即，若实矩阵 $\boldsymbol{A}$ 具有复数特征值

λ, 则 $\bar{\lambda}$ 也是 $\boldsymbol{A}$ 的特征值. 现在为了简单起见, 假定 $\boldsymbol{A}$ 只有一对复数特征值 λ_1, $\lambda_2=\bar{\lambda}_1$, 除此之外的特征值皆为实数.

$$\begin{matrix}\lambda_1=\sigma+j\omega \\ \lambda_2=\sigma-j\omega\end{matrix} \quad \sigma,\ \omega\in R \quad \omega\neq 0 \tag{22}$$

若设对应于 λ_1 的左特征向量为 $\boldsymbol{y}_1^T$, 则 $\bar{\boldsymbol{y}}^T$ 为对应于 λ_2 的左特征向量(试证明). 因此假定

$$\begin{matrix}\boldsymbol{y}_1^T=\boldsymbol{\eta}^T+j\boldsymbol{\xi}^T \\ \boldsymbol{y}_2^T=\boldsymbol{\eta}^T-j\boldsymbol{\xi}^T\end{matrix} \quad \eta、\xi\in R^n \tag{23}$$

[定理 1] 在 $\boldsymbol{u}=\boldsymbol{0}$ 的情况下, 无论起始状态选成什么样的实向量, 都不能只激励第 1 个模 $\tilde{x}_1$(第 2 个模 $\tilde{x}_2$).

[证明] 若是只激励第 1 个模, 则必须 $\tilde{x}_1(0)\neq 0$, $\tilde{x}_i(0)=0$, $i=2, \cdots, n$, 现假定 $\tilde{x}_2(0)=0$ 成立, 则

$$\tilde{x}_2(0)=\boldsymbol{y}_2^T\boldsymbol{x}(0)=\boldsymbol{\eta}^T\boldsymbol{x}(0)-j\boldsymbol{\xi}^T\boldsymbol{x}(0)=0 \tag{24}$$

因 $\boldsymbol{\eta}$, $\boldsymbol{\xi}$, $\boldsymbol{x}(0)$ 是实向量, 则 $\boldsymbol{\eta}^T\boldsymbol{x}(0)=\boldsymbol{\xi}^T\boldsymbol{x}(0)=0$. 因此

$$\tilde{x}_1(0)=\boldsymbol{y}_1^T\boldsymbol{x}(0)=\boldsymbol{\eta}^T\boldsymbol{x}(0)+j\boldsymbol{\xi}^T\boldsymbol{x}(0)=0 \tag{25}$$

下面性质是重要的.

[引理] $\boldsymbol{\eta}$ 和 $\boldsymbol{\xi}$ 是线性独立的.

[证明]

$$\begin{bmatrix}\boldsymbol{y}_1^T \\ \boldsymbol{y}_2^T\end{bmatrix}=\begin{bmatrix}1 & j \\ 1 & -j\end{bmatrix}\begin{bmatrix}\boldsymbol{\eta}^T \\ \boldsymbol{\xi}^T\end{bmatrix},$$

因

$$\det\begin{bmatrix}1 & j \\ 1 & -j\end{bmatrix}\neq 0.$$

显然, $\boldsymbol{\eta}$ 和 $\boldsymbol{\xi}$ 是线性独立的(参看 B-I-11 系 13).

利用该结果,在矩阵 $\boldsymbol{T}$ 中可以将 $\boldsymbol{y}_1^T$ 和 $\boldsymbol{y}_2^T$ 换成 $\boldsymbol{\xi}^T$ 和 $\boldsymbol{\eta}^T$

$$\boldsymbol{T}\triangleq\begin{bmatrix}\boldsymbol{y}_1^T \\ \boldsymbol{y}_2^T \\ \boldsymbol{y}_3^T \\ \vdots \\ \boldsymbol{y}_n^T\end{bmatrix}\longrightarrow\boldsymbol{T}_1\triangleq\begin{bmatrix}\boldsymbol{\xi}^T \\ \boldsymbol{\eta}^T \\ \boldsymbol{y}_3^T \\ \vdots \\ \boldsymbol{y}_n^T\end{bmatrix}\in R^{n\times n} \tag{26}$$

若由 T_1 代替 T 进行坐标变换, 则 $\tilde{\boldsymbol{A}}$ 变成如下形状(试证明)

$$\tilde{\boldsymbol{A}}=\boldsymbol{T}_1\boldsymbol{A}\boldsymbol{T}_1^{-1}=\left[\begin{array}{cc:c}\sigma & \omega & \multirow{2}{*}{$\boldsymbol{0}$} \\ -\omega & \sigma & \\ \hdashline \multicolumn{2}{c:}{\boldsymbol{0}} & \boldsymbol{J}_R\end{array}\right]\in R^{n\times n}$$

$$\boldsymbol{J}_R=\text{对于实特征值 } \lambda_3, \cdots, \lambda_n \text{ 的若当形} \tag{27}$$

这样, 在 $\boldsymbol{A}$ 具有复数特征值的情况下, 用实数范围内的变换, 对于 $\tilde{\boldsymbol{A}}$ 可以得到上述准对角形.

实现问题

在 A-I-7 中已证明过给定脉冲响应矩阵能用一般变系数的状态方程式和输出方程式实现的充分必要条件．这里我们将要讨论、满足什么样的条件才能用常系数的状态方程式和输出方程式实现．为此，首先脉冲响应矩阵必须是$(t-\tau)$的函数．

［定理 2］ 脉冲响应矩阵 $\boldsymbol{H}(t)$ 能用常系数的状态方程式和输出方程式实现的充分必要条件是，其拉氏变换 $\hat{\boldsymbol{H}}(s)$（即传递矩阵）是适宜矩阵．

［证明］ 必要性在(9)式中已经证过．为了证明其充分性，设 $\hat{\boldsymbol{H}}(s)$ 的第 i, j 个元素为[1)]

$$\hat{h}_{ij}(s)=\frac{b_{n-1}s^{n-1}+\cdots+b_1s+b_0}{s^n+a_{n-1}s^{n-1}+\cdots+a_1s+a_0}+d \tag{28}$$

若选取

$$\boldsymbol{A}_{ij}=\begin{bmatrix} 0 & 1 & & & \\ & \ddots & \ddots & & \boldsymbol{0} \\ & & \ddots & \ddots & \\ & \boldsymbol{0} & & \ddots & 1 \\ & & & & 0 \\ -a_0 & -a_1 & \cdots & & a_{n-1} \end{bmatrix},\qquad \boldsymbol{b}_{ij}=\begin{bmatrix}0\\ \vdots\\ \vdots\\ 0\\ 1\end{bmatrix} \tag{29}$$

$$\mathbf{c}_{ij}=[b_0b_1\cdots b_{n-1}]\quad d_{ij}=d$$

则

$$h_{ij}(s)=\boldsymbol{c}_{ij}(s\boldsymbol{I}-\boldsymbol{A}_{ij})^{-1}\boldsymbol{b}_{ij}+d_{ij} \tag{30}$$

成立（试证明）．因此，若设[2)]

$$\boldsymbol{A}\triangleq\boldsymbol{A}_1\dot{+}\boldsymbol{A}_2\dot{+}\cdots\dot{+}\boldsymbol{A}_m;$$

$$\boldsymbol{A}_i\triangleq\boldsymbol{A}_{i1}\dot{+}\cdots\dot{+}\boldsymbol{A}_{ir}\quad(1\leqslant i\leqslant m)$$

$$\boldsymbol{B}\triangleq\begin{bmatrix}\boldsymbol{B}_1\\ \vdots\\ \boldsymbol{B}_m\end{bmatrix};\quad \boldsymbol{B}_i=\boldsymbol{b}_{i1}\dot{+}\cdots\dot{+}\boldsymbol{b}_{ir}\quad(1\leqslant i\leqslant m) \tag{31}$$

$$\boldsymbol{C}\triangleq\boldsymbol{C}_1\dot{+}\cdots\dot{+}\boldsymbol{C}_m;\quad \boldsymbol{C}_i=[\mathbf{c}_{i1}\cdots\mathbf{c}_{ir}]\quad(1\leqslant i\leqslant m)$$

$$\boldsymbol{D}\triangleq[d_{ij}]$$

则

$$\boldsymbol{C}[s\boldsymbol{I}-\boldsymbol{A}]^{-1}\boldsymbol{B}+\boldsymbol{D}=\hat{\boldsymbol{H}}(s) \tag{32}$$

即可以作为

$$\boldsymbol{C}e^{\boldsymbol{A}t}\boldsymbol{B}+\boldsymbol{D}\delta(t)=\boldsymbol{H}(t) \tag{33}$$

$\boldsymbol{H}(t)$ 可以实现．

证明完毕

上面证明过程中得到的实现，一般不是最小实现．对于最小实现，将在关于可控性、可观测性部分说明．

1) $\hat{\boldsymbol{H}}(s)$ 是适宜的，与元素 $\hat{h}_{ij}(s)$ $(i,j=1,\cdots,n)$ 都是分母次数不小于分子次数的 s 的有理函数等价．因此，$\hat{h}_{ij}(s)$ 可以写成(28)式．

2) 这里为了方便起见，当矩阵 $\boldsymbol{S}$ 可以分割成

$$\boldsymbol{S}=\left[\begin{array}{c:c}\boldsymbol{S}_1 & 0\\ \hdashline 0 & \boldsymbol{S}_2\end{array}\right]$$

时，既使 $\boldsymbol{S}_1$ 及 $\boldsymbol{S}_2$, $\boldsymbol{S}$ 不是方阵，也写成 $\boldsymbol{S}=\boldsymbol{S}_1\dot{+}\boldsymbol{S}_2$．

A-I-11 可控性和可观测性

现在,我们来讨论线性常系数系统

$$\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{u} \tag{1a}$$

$$\boldsymbol{y}=\boldsymbol{C}\boldsymbol{x} \tag{1b}$$

的可控性和可观测性. A-I-8 中关于变系数系统讲过的一般结果, 当然在这里都适用. 因为已经知道可控性、可观测性与 $\boldsymbol{D}$ 无关,所以在(1b)式中略去了 $\boldsymbol{Du}$ 项.

因为线性常系数系统是定常的(见 A-I-9 中问题 3),则立即可得下面结果.

[定理 1] 线性常系数系统若在某个 t_0 时刻完全可控(完全可观测),则完全可控(完全可观测).

因此,以下都认为 $t_0=0$. 于是可以看到,若在 A-I-8 定理 1 中考虑 $\boldsymbol{\Phi}(t)=e^{At}$, 则系统(1a)式完全可控, $\Leftrightarrow$ 克兰姆矩阵 $\int_0^{t_f} e^{-At}\boldsymbol{B}\boldsymbol{B}^T e^{-A^Tt}dt$ 是正则的 $\Leftrightarrow e^{-At}\boldsymbol{B}$ 的行是线性独立的 $\Leftrightarrow \boldsymbol{P}^T e^{-At}\boldsymbol{B}=\boldsymbol{0}\ {}^{\forall}t\in R$ 成立只限于 $\boldsymbol{P}=\boldsymbol{0}$.

假定对于某个 $\boldsymbol{P}\in R^n$

$$\boldsymbol{P}^T e^{-At}\boldsymbol{B}=\boldsymbol{0}\quad {}^{\forall}t\in R \tag{2}$$

成立. 将上式两边对 t 微分后, 令 $t=0$, 则得到

$$\boldsymbol{P}^T\boldsymbol{B}=\boldsymbol{0},\ \boldsymbol{P}^T\boldsymbol{A}\boldsymbol{B}=\boldsymbol{0},\ \cdots\cdots,\ \boldsymbol{P}^T\boldsymbol{A}^k\boldsymbol{B}=\boldsymbol{0},\ \cdots\cdots$$

因此,若 $e^{-At}\boldsymbol{B}$ 的行不是线性独立的,即,对于某个 $\boldsymbol{P}\neq\boldsymbol{0}$, 若(2)式成立,则

$$\boldsymbol{P}^T[\boldsymbol{B}\ \ \boldsymbol{A}\boldsymbol{B}\cdots\boldsymbol{A}^{n-1}\boldsymbol{B}]=\boldsymbol{0} \tag{3}$$

成立,矩阵

$$\boldsymbol{Q}_n\triangleq[\boldsymbol{B}\ \ \boldsymbol{A}\boldsymbol{B}\cdots\boldsymbol{A}^{n-1}\boldsymbol{B}] \tag{4}$$

的行不是线性独立的. 反之, 若假定 $\boldsymbol{Q}$ 的行不是线性独立的, 则对于某个 $\boldsymbol{P}\neq 0$, (3)式成立. e^{At} 可以用 $\boldsymbol{I}$, $\boldsymbol{A}$, $\cdots$, $\boldsymbol{A}^{n-1}$ 的线性组合表示(见 A-I-9 中(21)式), 因此(2)式成立, 即 $e^{At}\boldsymbol{B}$ 的行不是线性独立的. 由以上得到下面结果[1].

[定理 2] 线性常系数系统((1a)式)完全可控的充分必要条件是

$$\operatorname{rank}\boldsymbol{Q}_n=\operatorname{rank}[\boldsymbol{B}\ \ \boldsymbol{A}\boldsymbol{B}\cdots\boldsymbol{A}^{n-1}\boldsymbol{B}]=n \tag{5}$$

根据完全相同的推导,可得下面定理.

[定理 3] 线性常系数系统((1)式)完全可观测的充分必要条件是

$$\operatorname{rank}\boldsymbol{R}_n=\operatorname{rank}[\boldsymbol{C}^T\boldsymbol{A}^T\boldsymbol{C}^T\cdots(\boldsymbol{A}^T)^{n-1}\boldsymbol{C}^T]^T=n \tag{6}$$

然而,若 $\boldsymbol{A}^k\boldsymbol{B}$ 是 $\boldsymbol{B}$, $\boldsymbol{A}\boldsymbol{B}$, $\cdots$, $\boldsymbol{A}^{k-1}\boldsymbol{B}$ 的线性组合, 则 $\boldsymbol{A}^l\boldsymbol{B}\,(l\geqslant k)$ 也可以用 $\boldsymbol{B}$, $\boldsymbol{A}\boldsymbol{B}$, $\cdots$, $\boldsymbol{A}^{k-1}\boldsymbol{B}$ 的线性组合表示(可以和 A-I-9 中(27)式同样证明). 因此,有下面系.

[系 1] 若定义矩阵 $\boldsymbol{Q}_k$ 为

$$\boldsymbol{Q}_k\triangleq[\boldsymbol{B}\ \ \boldsymbol{A}\boldsymbol{B}\cdots\boldsymbol{A}^{k-1}\boldsymbol{B}]\quad k=1,\ 2,\ \cdots\cdots \tag{7}$$

设 α 为使

1) 显然,定理 2 也可以看成 A-I-8 中系 2 的特殊情况.

$$\operatorname{rank} \boldsymbol{Q}_\alpha = \operatorname{rank} \boldsymbol{Q}_{\alpha+1}$$

成立的最小自然数，则完全可控的充分必要条件是

$$\operatorname{rank} \boldsymbol{Q}_\alpha = n \tag{8}$$

其中 α 称为可控性指标(controllability index)。

关于可观测性同样有下面性质.

［系 2］ 若定义矩阵 $\boldsymbol{R}_k$ 为

$$\boldsymbol{R}_k \triangleq [\boldsymbol{C}^T \boldsymbol{A}^T \boldsymbol{C}^T \cdots (\boldsymbol{A}^T)^{k-1} \boldsymbol{C}^T]^T \tag{9}$$

设 β 为使

$$\operatorname{rank} \boldsymbol{R}_\beta = \operatorname{rank} \boldsymbol{R}_{\beta+1}$$

成立的最小自然数，则完全可观测的充分必要条件是

$$\operatorname{rank} \boldsymbol{R}_\beta = n \tag{10}$$

其中 β 称为可观测性指标(observability index).

而且，有下列性质.

［系 3］ 系统((1)式)完全可控完全可观测的充分必要条件是，当 α, β 按系 1，2 定义时，

$$\operatorname{rank} \boldsymbol{R}_\beta \boldsymbol{Q}_\alpha = n \tag{11}$$

成立.

［证明］ 若(11)式成立，则[1]

$$\operatorname{rank} \boldsymbol{R}_\beta \geqslant n, \ \operatorname{rank} \boldsymbol{Q}_\alpha \geqslant n.$$

因 $\boldsymbol{R}_\beta$ 有 n 个列，$\boldsymbol{Q}_\alpha$ 有 n 个行，二者的秩都不超过 n，则(8)，(10)式成立.

反之，若(8)，(10)式成立，则 $\boldsymbol{R}_\beta$ 中含有 n 个线性独立的行，$\boldsymbol{Q}_\alpha$ 中含有 n 个线性独立的列. 在 $\boldsymbol{R}_\beta$, $\boldsymbol{Q}_\alpha$ 中留下这些线性独立的行或列，去掉其它部分，设所得到的子矩阵分别为 $\widetilde{\boldsymbol{R}}_\beta$, $\widetilde{\boldsymbol{Q}}_\alpha$，二者皆为 $n\times n$ 阶正则矩阵. 因此，$\det(\widetilde{\boldsymbol{R}}_\beta \widetilde{\boldsymbol{Q}}_\alpha) \neq 0$. 因为显然 $\widetilde{\boldsymbol{R}}_\beta \widetilde{\boldsymbol{Q}}_\alpha$ 是 $\boldsymbol{R}_\beta \boldsymbol{Q}_\alpha$ 的子矩阵，则 $\boldsymbol{R}_\beta \boldsymbol{Q}_\alpha$ 中存在不为 0 的 n 级子式，$\operatorname{rank} \boldsymbol{R}_\beta \boldsymbol{Q}_\alpha \geqslant n$. 根据(8)，(10)式，因不会超过 n[2]，故得(11)式.

证明完毕

(注意)

“(8)，(10)式 $\Leftrightarrow$ (11)式”可以直接用西勒维斯特不等式 (B-I-3 中(23)式)证明.

在 $m=r$，即输出数与输入数相等的情况下，也可以定义积 $\boldsymbol{Q}_n \boldsymbol{R}_n$. 若 $\operatorname{rank}(\boldsymbol{Q}_n \boldsymbol{R}_n) = n$，则系统((1)式)是可控且可观测的. 但是，一般其逆不成立. 例如，我们来考察下列 2 输入 2 输出系统

$$\dot{\boldsymbol{x}} = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ 0 & 1 & -1 \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \boldsymbol{u} \tag{12a}$$

$$\boldsymbol{y} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \boldsymbol{x} \tag{12b}$$

1) 一般，对于两个矩阵 $\boldsymbol{A}=m\times p$, $\boldsymbol{B}=p\times n$ 有下列关系

$$\operatorname{rank}(\boldsymbol{AB}) \leqslant \min(\operatorname{rank}\boldsymbol{A}, \operatorname{rank}\boldsymbol{B}).$$

该不等式可由秩的定义及 P-I-2 中(15)式直接得出.

2) 同上.

对于该系统，因

$$\boldsymbol{Q}_n=\begin{bmatrix}1&0&1&0&2&0\\0&1&1&0&2&0\\0&1&0&0&1&0\end{bmatrix} \tag{13}$$

$$\boldsymbol{R}_n=\begin{bmatrix}1&-1&0\\0&0&1\\0&0&0\\0&1&-1\\0&0&0\\1&0&0\end{bmatrix} \tag{14}$$

rank $\boldsymbol{Q}_n$=rank $\boldsymbol{R}_n=3(=n)$，则该系统可控且可观测．但是

$$\text{rank}\ \boldsymbol{Q}_n\boldsymbol{R}_n=\text{rank}\begin{bmatrix}1&-1&0\\0&0&1\\0&0&1\end{bmatrix}=2<n \tag{15}$$

当然，若 $m=r=1$，则 rank $\boldsymbol{Q}_n$，$\boldsymbol{R}_n=n$ 是可控且可观测的充分必要条件．

矩阵 $\boldsymbol{Q}_n$，$\boldsymbol{R}_n$ 分别称为可控性矩阵(controllability matrix)，可观测性矩阵(observability matrix)．以下广泛使用这个术语，设 α，β 为系1，2中定义的自然数时，$\boldsymbol{Q}_k(k\geqslant\alpha)$ 均称为可控性矩阵，$\boldsymbol{R}_l(l\geqslant\beta)$ 均称为可观测性矩阵．

积 $\boldsymbol{R}_\beta\boldsymbol{Q}_\alpha$ 是什么样的矩阵呢？显而易见

$$\boldsymbol{R}_\beta\boldsymbol{Q}_\alpha=\begin{bmatrix}\boldsymbol{CB}&\boldsymbol{CAB}&\cdots&\boldsymbol{CA}^{\alpha-1}\boldsymbol{B}\\\boldsymbol{CAB}&\boldsymbol{CA}^2\boldsymbol{B}&\cdots&\boldsymbol{CA}^{\alpha}\boldsymbol{B}\\\vdots&\vdots&&\vdots\\\boldsymbol{CA}^{\beta-1}\boldsymbol{B}&\boldsymbol{CA}^{\beta}\boldsymbol{B}&\cdots&\boldsymbol{CA}^{\alpha+\beta-1}\boldsymbol{B}\end{bmatrix}=m\beta\times r\alpha \tag{16}$$

它是以 $\boldsymbol{CA}^{i+j-2}\boldsymbol{B}$ 为第$(i,\ j)$个子块的分块矩阵．当给出脉冲响应矩阵 $\boldsymbol{H}(t)$ 时，第$(i,\ j)$个子块可以由下式

$$\text{第}(i,\ j)\text{个子块}=\left[\frac{d^{i+j-2}}{dt^{i+j-2}}\boldsymbol{H}(t)\right]_{t=0} \tag{17}$$

求出．

在这里，我们来考察一下 α，β 值的范围．首先，设 $\boldsymbol{A}$ 的最小多项式的次数为 ν，则显然 α，$\beta\leqslant\nu$．

因 $\boldsymbol{B}=n\times r$，则 $\boldsymbol{Q}_k=n\times kr$．因此，为使(8)式成立，则必须

$$\alpha\geqslant\frac{n}{r}. \tag{18}$$

现假定

$$\text{rank}\ \boldsymbol{B}=r. \tag{19}$$

在 $\boldsymbol{A}^i\boldsymbol{B}(i=1,\ 2,\ \cdots,\ \alpha-1)$ 的列中，至少有一个不能用 $\boldsymbol{B}$，$\boldsymbol{AB}$，…，$\boldsymbol{A}^{i-1}\boldsymbol{B}$ 的列线性组合表示(为何)．因此

$$r+\alpha-1\leqslant n \tag{20}$$

由以上得

$$\frac{n}{r} \leqslant \alpha \leqslant \min(\nu,\ n-r+1) \tag{21}$$

同样，对于 β 下式成立

$$\frac{n}{m} \leqslant \beta \leqslant \min(\nu,\ n-m-1). \tag{22}$$

［问题 1］ 试证明当且仅当线性常系数系统完全可控(完全可观测)时，该系统是一致可控(一致可观测)的.

现在，我们设 $\boldsymbol{A}$ 的若当标准形 $\boldsymbol{J}$ 为 $\boldsymbol{J}=\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$，定义 $\widetilde{\boldsymbol{B}} \triangleq \boldsymbol{T}\boldsymbol{B}$，将 $\widetilde{\boldsymbol{B}}$ 对应于 $\boldsymbol{J}$ 的子块进行分割，即

$$\boldsymbol{J}=\begin{bmatrix} \boldsymbol{J}_1 & & \boldsymbol{0} \\ & \ddots & \\ \boldsymbol{0} & & \boldsymbol{J}_\sigma \end{bmatrix} \quad \widetilde{\boldsymbol{B}}=\begin{bmatrix} \widetilde{\boldsymbol{B}}_1 \\ \cdots \\ \vdots \\ \cdots \\ \widetilde{\boldsymbol{B}}_\sigma \end{bmatrix} \quad \widetilde{\boldsymbol{B}}_j = m_j \times r \quad j=1,\ \cdots,\ \sigma \tag{23}$$

$$\boldsymbol{J}_j=\begin{bmatrix} \boldsymbol{J}_{j1} & & \boldsymbol{0} \\ & \ddots & \\ \boldsymbol{0} & & \boldsymbol{J}_{j\alpha_j} \end{bmatrix} \quad \widetilde{\boldsymbol{B}}_j=\begin{bmatrix} \widetilde{\boldsymbol{B}}_{j1} \\ \cdots \\ \vdots \\ \cdots \\ \widetilde{\boldsymbol{B}}_{j\alpha_j} \end{bmatrix} \quad \widetilde{\boldsymbol{B}}_{jk} = n_{jk} \times r \quad \begin{matrix} k=1,\ \cdots,\ \alpha_j \\ j=1,\ \cdots,\ \sigma \end{matrix} \tag{24}$$

而且，设 $\widetilde{\boldsymbol{B}}_{jk}$ 的最下一行为 $\boldsymbol{\beta}_{jk}$.

［定理 4］ 线性常系数系统((1a)式)完全可控的充分必要条件是

$$\operatorname{rank}\begin{bmatrix} \boldsymbol{\beta}_{j1} \\ \cdots \\ \boldsymbol{\beta}_{j\alpha_j} \end{bmatrix} = \alpha_j \quad j=1,\ \cdots,\ \sigma \tag{25}$$

（证明） 由上面 $\boldsymbol{J}$, $\widetilde{\boldsymbol{B}}$ 的表示式得

$$e^{J}\widetilde{\boldsymbol{B}}=\begin{bmatrix} e^{J_1 t}\widetilde{\boldsymbol{B}}_1 \\ \cdots\cdots\cdots \\ \vdots \\ \cdots\cdots\cdots \\ e^{J_\sigma t}\widetilde{\boldsymbol{B}}_\sigma \end{bmatrix} \tag{24}$$

根据 B-I-13(16)式，$e^{J_j t}\widetilde{\boldsymbol{B}}_j$ 的各元素是 $e^{\lambda_j t}$, $te^{\lambda_j t}$, $\cdots$, $t^{(d_j-1)}e^{\lambda_j t}$ 的线性组合．一般，若 λ 的值相异，则对于任意 C_{jk} $\sum_{k=1}^{d_j} C_{jk}t^{(k-1)}e^{\lambda_j t}$ 是线性独立的．亦即，若 $\lambda_j(j=1,\ \cdots,\ \sigma)$ 相异，则

$$\sum_{j=1}^{\sigma} P_j \sum_{k=1}^{d_j} C_{jk}t^{(k-1)}e^{\lambda_j t}=0 \quad {}^\forall t\in R \Leftrightarrow P_j=0 \quad j=1,\ \cdots,\ \sigma \tag{25}$$

根据该性质，对于 $\boldsymbol{P} \triangleq [\boldsymbol{P}_1^T \cdots \boldsymbol{P}_\sigma^T]^T$ ($\boldsymbol{P}_j$ 是 m_j 维向量)，可得到

$$\boldsymbol{P}^T e^{Jt}\widetilde{\boldsymbol{B}}=\sum_{j=1}^{\sigma} \boldsymbol{P}_j^T e^{J_j t}\widetilde{\boldsymbol{B}}_j=\boldsymbol{0} \quad {}^\forall t\in R \quad \Leftrightarrow \boldsymbol{P}_j^T e^{J_j t}\widetilde{\boldsymbol{B}}_j=\boldsymbol{0} \quad j=1,\ \cdots,\ \sigma \quad {}^\forall t\in R \tag{26}$$

即，完全可控的充分必要条件是，对于各个 $j=1,\ \cdots,\ \sigma$, $e^{J_j t}\widetilde{\boldsymbol{B}}_j$ 的各行是线性独立的.

其次，再对

$$e^{J_jt}\widetilde{\boldsymbol{B}}_j=\begin{bmatrix} e^{J_{j1}t}\widetilde{\boldsymbol{B}}_{j1} \\ \cdots\cdots\cdots \\ \vdots \\ \cdots\cdots\cdots \\ e^{J_{j\alpha_j}t}\widetilde{\boldsymbol{B}}_{j\alpha_j} \end{bmatrix} \tag{27}$$

使用 B-I-13 中(16)式，可见 $e^{J_{jk}t}\widetilde{\boldsymbol{B}}_{jk}$ 的最下一行等于 $e^{\lambda_jt}\boldsymbol{\beta}_{jk}$。因此，若(25)式不成立，假定在某个 j 时 $\boldsymbol{\beta}_{j1}, \cdots, \boldsymbol{\beta}_{j\alpha_j}$ 线性相关，则 $e^{J_jt}\widetilde{\boldsymbol{B}}_j$ 的行线性相关。即，(25)式是使 $e^{J_jt}\widetilde{\boldsymbol{B}}_j$ 的行线性独立的必要条件。

下面来证明(25)式也是充分条件。为便于理解，设

$$\alpha_j=2,\ n_{j1}=n_{j2}+1 \tag{28}$$

一般情况也可以同样证明。对于 $\boldsymbol{P}\triangleq[\boldsymbol{P}_1^T\boldsymbol{P}_2^T]^T$ ($\boldsymbol{P}_k$ 是 n_{jk} 维向量)，在(25)式基础上，假定

$$\boldsymbol{P}^Te^{J_jt}\widetilde{\boldsymbol{B}}_j=\boldsymbol{P}_1^Te^{J_{j1}t}\widetilde{\boldsymbol{B}}_{j1}+\boldsymbol{P}_2^Te^{J_{j2}t}\widetilde{\boldsymbol{B}}_{j2}=\boldsymbol{0} \tag{29}$$

成立。因为左边是 $e^{\lambda_jt}, te^{\lambda_jt}, \cdots, t^{(n_{j1}-1)}e^{\lambda_jt}$ 的线性组合，为使其等于 $\boldsymbol{0}$，则 $t^le^{\lambda_jt}$ ($l=0, 1, \cdots, n_{j1}-1$) 的系数必须均为 0。首先，$t^{(n_{j1}-1)}e^{\lambda_jt}$ 的系数等于 $P_{11}\boldsymbol{\beta}_{j1}$。因 $\boldsymbol{\beta}_{j1}\neq\boldsymbol{0}$(否则(25)式不成立)，则 $P_{11}=0$。其次，考虑到 $P_{11}=0$，则 $t^{(n_{j1}-2)}e^{\lambda_jt}$，即 $t^{(n_{j2}-1)}e^{\lambda_jt}$ 的系数变成 $(P_{12}\boldsymbol{\beta}_{j1}+P_{21}\boldsymbol{\beta}_{j2})$。若(25)式成立，则只有在 $P_{12}=P_{21}=0$ 时才有 $P_{12}\boldsymbol{\beta}_{j1}+P_{21}\boldsymbol{\beta}_{j2}=\boldsymbol{0}$。此外，若逐个考察一下 $t^{(n_{j1}-3)}e^{\lambda_jt}$……的系数，可以证明 $\boldsymbol{P}=\boldsymbol{0}$，即 $e^{J_jt}\widetilde{\boldsymbol{B}}_j$ 的行是线性独立的。

［系 4］ 设 $\boldsymbol{A}$ 的特征值均相异，可按 $\boldsymbol{\Lambda}=\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$ 化成对角形，则完全可控的充分必要条件是，在 $\boldsymbol{T}^{-1}\boldsymbol{B}$ 中不存在等于零向量的行。

［系 5］ 为使 $r=1$(即单输入)系统是可控的，则 $\boldsymbol{A}$ 的特征多项式与最小多项式必须一致。

(问题 2) 试证明系 4, 5 (系 5 也可以用(18)式证明)。

(问题 3) 对于可观测性，试推导出与上述相当的结果。

线性常系数系统的典范结构(Canonical Structure)

这里将要叙述大家所熟悉的用下列标准形表示的线性常系数系统的卡尔曼典范结构分解定理

$$\begin{aligned} \dot{\boldsymbol{x}}&=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{u} \quad \boldsymbol{A}\in R^{n\times n},\ \boldsymbol{B}\in R^{n\times r} \\ \boldsymbol{y}&=\boldsymbol{C}\boldsymbol{x}+\boldsymbol{D}\boldsymbol{u} \quad \boldsymbol{C}\in R^{m\times n},\ \boldsymbol{D}\in R^{m\times r} \end{aligned} \tag{30}$$

该定理内容大概如下。在研究系统的可控性和可观测性时，若选择适当的变换矩阵 $\boldsymbol{T}\in R^{n\times n}$，则变换后(即关于 $\boldsymbol{T}$ 代数等价)的系统，可以用具有下列性质的四个“子系统”的组合表示，即可控可观测、可控不可观测、不可控可观测和不可控不可观测。

该典范结构定理最重要的意义是，它表明系统(30)的传递函数仅能表示出上述四个子系统中可控可观测的那个子系统的特性，其它子系统的特性与传递函数完全无关。也就是说，传递函数仅包含系统动力学的一部分信息。

作为该结果的直接应用，用标准形实现给定传递函数时的最小实现性质是很清楚的。回忆一下上述内容，当实现传递函数的某个标准形是可控可观测时，称它为给定传递函数的最紧实现，即称为最小实现。

在这里为了节省篇幅，在叙述下述内容之前先规定一个符号。因为用(30)式标准形

表示的系统完全由矩阵 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$ 规定，所以该系统我们以后用 $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D})$ 表示，特别是当 $\boldsymbol{D}=\boldsymbol{0}$ 时，写成 $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$.

［定理 5］ 设 $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ 的可控性指标为 α, $\operatorname{rank}\boldsymbol{Q}_\alpha=q^c(\leqslant n)$. 这时有某个正则矩阵 $\boldsymbol{T}\in R^{n\times n}$ 存在，对于 $\boldsymbol{T}$, $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ 和满足下列条件 (i), (ii) 的 $(\boldsymbol{A}^c, \boldsymbol{B}^c, \boldsymbol{C}^c)$ 代数等价.

(i)

$$\left.\begin{aligned}
\boldsymbol{A}^c &\triangleq \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\begin{bmatrix}\overbrace{\boldsymbol{A}_{11}^c}^{q^c} & \overbrace{\boldsymbol{A}_{12}^c}^{n-q^c}\\ \boldsymbol{0} & \boldsymbol{A}_{22}\end{bmatrix}\begin{matrix}\}q^c\\ \}n-q^c\end{matrix}\\
\boldsymbol{B}^c &\triangleq \boldsymbol{T}^{-1}\boldsymbol{B}=\begin{bmatrix}\boldsymbol{B}_1^c\\ \boldsymbol{0}\end{bmatrix}\begin{matrix}\}q^c\\ \}n-q^c\end{matrix}\\
\boldsymbol{C}^c &\triangleq \boldsymbol{C}\boldsymbol{T}=[\overbrace{\boldsymbol{C}_1^c}^{q^c}\ \overbrace{\boldsymbol{C}_2^c}^{n-q^c}]
\end{aligned}\right\}\tag{31}$$

(ii) 由子矩阵组成的 q^c 维标准形 $(\boldsymbol{A}_{11}^c, \boldsymbol{B}_1^c, \boldsymbol{C}_1^c)$ 是可控的，且其传递函数与 $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ 一致.

（证明） (i) 因 $\operatorname{rank}\boldsymbol{Q}_\alpha=q^c$, 则在 $\boldsymbol{Q}_\alpha$ 的 αr 个列向量中有 q^c 个线性独立组 (B-I-3 中定理 2). 将此归纳，设 $\boldsymbol{T}_1\in R^{n\times q^c}$, 另外适当确定 $\boldsymbol{T}_2\in R^{n\times(n-q^c)}$, 使 $\boldsymbol{T}\triangleq[\boldsymbol{T}_1, \boldsymbol{T}_2]\in R^{n\times n}$ 为正则矩阵，而且写成

$$\boldsymbol{T}^{-1}\triangleq\begin{bmatrix}\tilde{\boldsymbol{T}}_1\\ \tilde{\boldsymbol{T}}_2\end{bmatrix}\begin{matrix}\}q^c\\ \}n-q^c.\end{matrix}$$

这里由 $\boldsymbol{T}$ 变换 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, 首先，因

$$\boldsymbol{A}^c\triangleq\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\begin{bmatrix}\tilde{\boldsymbol{T}}_1\\ \tilde{\boldsymbol{T}}_2\end{bmatrix}[\boldsymbol{A}\boldsymbol{T}_1\boldsymbol{A}\boldsymbol{T}_2]\tag{32}$$

$\boldsymbol{A}\boldsymbol{T}_1$ 的各列包含于 $\boldsymbol{Q}_{\alpha+1}$ 的列，$\boldsymbol{Q}_{\alpha+1}$ 的列可以用 $\boldsymbol{T}_1$ 线性表出（因根据假定 $\operatorname{rank}\boldsymbol{Q}_\alpha=\operatorname{rank}\boldsymbol{Q}_{\alpha+1}$），则对于某个 $\boldsymbol{A}_{11}^c\in R^{q^c\times q^c}$ 可以写成 $\boldsymbol{A}\boldsymbol{T}_1=\boldsymbol{T}_1\boldsymbol{A}_{11}^c$. 考虑到这一点，(32) 式可表示成

$$\begin{aligned}
\boldsymbol{A}^c&=\begin{bmatrix}\tilde{\boldsymbol{T}}_1\\ \tilde{\boldsymbol{T}}_2\end{bmatrix}[\boldsymbol{T}_1\boldsymbol{A}_{11}^c\boldsymbol{A}\boldsymbol{T}_2]=\begin{bmatrix}\boldsymbol{A}_{11}^c & \tilde{\boldsymbol{T}}_1\boldsymbol{A}\boldsymbol{T}_2\\ \boldsymbol{0} & \tilde{\boldsymbol{T}}_2\boldsymbol{A}\boldsymbol{T}_2\end{bmatrix}\\
&\triangleq\begin{bmatrix}\boldsymbol{A}_{11}^c & \boldsymbol{A}_{12}^c\\ \boldsymbol{0} & \boldsymbol{A}_{22}^c\end{bmatrix}
\end{aligned}$$

同样，因 $\boldsymbol{B}$ 含于 $\boldsymbol{Q}_\alpha$, 则对于某个 $\boldsymbol{B}_1\in R^{q^c\times r}$ 可以表示成 $\boldsymbol{B}=\boldsymbol{T}_1\boldsymbol{B}_1^c$.

$$\boldsymbol{B}^c\triangleq\boldsymbol{T}^{-1}\boldsymbol{B}=\begin{bmatrix}\tilde{\boldsymbol{T}}_1\\ \tilde{\boldsymbol{T}}_2\end{bmatrix}\boldsymbol{T}_1\boldsymbol{B}_1^c=\begin{bmatrix}\boldsymbol{B}_1^c\\ \boldsymbol{0}\end{bmatrix}$$

$\boldsymbol{C}^c$ 的表示式只不过是把 $\boldsymbol{C}^c\triangleq\boldsymbol{C}\boldsymbol{T}$ 分写成两个子矩阵.

(ii) 由 (31) 式可见，$(\boldsymbol{A}^c, \boldsymbol{B}^c, \boldsymbol{C}^c)$ 的可控性矩阵 $\boldsymbol{Q}_\alpha^c$ 有下列形状

$$\begin{aligned}
\boldsymbol{Q}_\alpha^c&\triangleq[\boldsymbol{B}^c\boldsymbol{A}^c\boldsymbol{B}^c\cdots\boldsymbol{A}^{c\alpha-1}\boldsymbol{B}^c]\\
&=\begin{bmatrix}\boldsymbol{B}_1^c & \boldsymbol{A}_{11}^c\boldsymbol{B}_1^c\cdots & \boldsymbol{A}_{11}^{c\alpha-1}\boldsymbol{B}_1^c\\ \boldsymbol{0} & \boldsymbol{0}\quad\cdots & \boldsymbol{0}\end{bmatrix}\begin{matrix}\}q^c\\ \}n-q^c.\end{matrix}
\end{aligned}\tag{33}$$

而且，因

$$\boldsymbol{Q}_\alpha^c=\boldsymbol{T}^{-1}\boldsymbol{Q}_\alpha\tag{34}$$

则容易证明 $\operatorname{rank}\boldsymbol{Q}_\alpha^c=\operatorname{rank}\boldsymbol{Q}_\alpha=q^c$ 成立*. 因此，根据 (33) 式，$(\boldsymbol{A}_{11}^c, \boldsymbol{B}_1^c, \boldsymbol{C}_1^c)$ 可控.

* 原文误印成 $\operatorname{rank}\boldsymbol{Q}_\alpha^c=\operatorname{rank}\boldsymbol{Q}_\alpha^c=q^c$ 成立. ——译者

它和$(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$具有同样的传递函数，例如可直接证明如下.

根据分块矩阵运算 (B-I-1 中系 3),

$$[s\boldsymbol{I}-\boldsymbol{A}^c]^{-1}=\begin{bmatrix} [s\boldsymbol{I}-\boldsymbol{A}_{11}^c]^{-1} & [s\boldsymbol{I}-\boldsymbol{A}_{11}^c]^{-1}\boldsymbol{A}_{12}^c[s\boldsymbol{I}-\boldsymbol{A}_{22}^c]^{-1} \\ \boldsymbol{0} & [s\boldsymbol{I}-\boldsymbol{A}_{22}^c]^{-1} \end{bmatrix}$$

而且用该式得

$$\begin{aligned} \boldsymbol{C}[s\boldsymbol{I}-\boldsymbol{A}]^{-1}\boldsymbol{B} &= \boldsymbol{C}\boldsymbol{T}[s\boldsymbol{I}-\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}]^{-1}\boldsymbol{T}^{-1}\boldsymbol{B} \\ &= \boldsymbol{C}^c[s\boldsymbol{I}-\boldsymbol{A}^c]^{-1}\boldsymbol{B}^c = \boldsymbol{C}_1^c[s\boldsymbol{I}-\boldsymbol{A}_{11}^c]^{-1}\boldsymbol{B}_1^c \end{aligned}$$

即$(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$和$(\boldsymbol{A}_{11}^c, \boldsymbol{B}_1^c, \boldsymbol{C}_1^c)$的传递函数一致.

证明完毕

(注意)

"(i) 为了具体地求变换矩阵 $\boldsymbol{T}$, 由(33)、(34)式可见, 对 $\boldsymbol{Q}_\alpha$ 进行行变换即可. 即, 将 $\boldsymbol{Q}_\alpha$ 行变换成

$$\boldsymbol{Q}_\alpha \longrightarrow \boldsymbol{Q}_\alpha^c = \begin{array}{l} \begin{bmatrix} \boldsymbol{Q}_{\alpha 1}^c \\ \boldsymbol{0} \end{bmatrix} \begin{array}{l} \}q^c \\ \}n-q^c \end{array} \end{array} \qquad \operatorname{rank}\boldsymbol{Q}_{\alpha 1}^c = q^c$$

的行变换矩阵相当于 $\boldsymbol{T}^{-1}$. 当然这样的 $\boldsymbol{T}^{-1}$ 不是唯一确定的, 选取其中的那一个都可以. 为了确定 $\boldsymbol{T}^{-1}$, 可以利用 B-I-2 中的方法. 考虑扩大的 $n\times(\alpha r+n)$ 矩阵$[\boldsymbol{Q}_\alpha \boldsymbol{I}_n]$,

$$[\boldsymbol{Q}_\alpha \boldsymbol{I}_n] \longrightarrow \left[\begin{array}{c|c} \boldsymbol{Q}_{\alpha 1}^c & \boldsymbol{T}^{-1} \\ \boldsymbol{0} & \end{array}\right]$$

注视 $\boldsymbol{Q}_\alpha$ 部分, 假定将其行变换成上面形状, 这时右半个子块将给出 $\boldsymbol{T}^{-1}$.

(ii) 一般, 两个标准形$(\boldsymbol{A}_1, \boldsymbol{B}_1, \boldsymbol{C}_1, \boldsymbol{D}_1)$和$(\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{C}_2, \boldsymbol{D}_2)$(它们不一定有同样的维数)的传递函数相等意味着, 其零状态响应(对于任意输入)完全一致.

因此, 这时$(\boldsymbol{A}_1, \boldsymbol{B}_1, \boldsymbol{C}_1, \boldsymbol{D}_1)$和$(\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{C}_2, \boldsymbol{D}_2)$称为是零状态等价 (zero-state equivalent)的. 利用该术语则定理 5 可说成$(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$和$(\boldsymbol{A}_{11}^c, \boldsymbol{B}_1^c, \boldsymbol{C}_1^c)$是零状态等价的."

此外, 对于$(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$的可观测性, 也可以得到和定理 5 相对应的结果.

[定理 6] 设$(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$的可观测性指标为 β, $\operatorname{rank}\boldsymbol{R}_\beta = q^0 (\leqslant n)$. 这时有某个正则矩阵 $\boldsymbol{T}\in R^{n\times n}$ 存在, 对于 $\boldsymbol{T}$, $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$和满足下列条件(i), (ii)的$(\boldsymbol{A}^0, \boldsymbol{B}^0, \boldsymbol{C}^0)$代数等价.

(i)

$$\left.\begin{aligned} \boldsymbol{A}^0 &\triangleq \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \begin{array}{c} \begin{array}{cc} \overbrace{\phantom{A_{11}^0}}^{q^0} & \overbrace{\phantom{A_{22}^0}}^{n-q^0} \end{array} \\ \begin{bmatrix} \boldsymbol{A}_{11}^0 & \boldsymbol{0} \\ \boldsymbol{A}_{21}^0 & \boldsymbol{A}_{22}^0 \end{bmatrix} \end{array} \begin{array}{l} \}q^0 \\ \}n-q^0 \end{array} \\ \boldsymbol{B}^0 &\triangleq \boldsymbol{T}^{-1}\boldsymbol{B} = \begin{bmatrix} \boldsymbol{B}_1^0 \\ \boldsymbol{B}_2^0 \end{bmatrix} \begin{array}{l} \}q^0 \\ \}n-q^0 \end{array} \\ \boldsymbol{C}^0 &\triangleq \boldsymbol{C}\boldsymbol{T} = [\overbrace{\boldsymbol{C}_1^0}^{q^0} \;\; \overbrace{\boldsymbol{0}}^{n-q^0}] \end{aligned}\right\} \tag{35}$$

(ii) 由子矩阵组成的 q^0 维标准形$(\boldsymbol{A}_{11}^0, \boldsymbol{B}_1^0, \boldsymbol{C}^0)$是可观测的, 且其传递函数和$(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$一致.

定理 6 的证明大体上与定理 5 一样.

(证明) 对$(\boldsymbol{A}^T, \boldsymbol{C}^T, \boldsymbol{B}^T)$用定理5，则有某个正则矩阵$\boldsymbol{U}\in R^{n\times n}$存在，使得

$$\boldsymbol{A}^{0T}\triangleq \boldsymbol{U}^{-1}\boldsymbol{A}^T\boldsymbol{U}=\begin{bmatrix}\boldsymbol{A}_{11}^{0T} & \boldsymbol{A}_{21}^{0T}\\ \boldsymbol{0} & \boldsymbol{A}_{22}^{0T}\end{bmatrix}$$

$$\boldsymbol{C}^{0T}\triangleq \boldsymbol{U}^{-1}\boldsymbol{C}^T=\begin{bmatrix}\boldsymbol{C}_1^{0T}\\ \boldsymbol{0}\end{bmatrix}$$

$$\boldsymbol{B}^{0T}=\boldsymbol{B}^T\boldsymbol{U}=[\,\boldsymbol{B}_1^{0T}\ \ \boldsymbol{B}_2^{0T}\,]$$

因此，作为(35)式的变换矩阵可以用$\boldsymbol{T}\triangleq(\boldsymbol{U}^{-1})^T$。

证明完毕

(注意)

“考虑到$(\boldsymbol{A}^0, \boldsymbol{B}^0, \boldsymbol{C}^0)$的可观测性矩阵$\boldsymbol{R}_\beta^0$

$$\boldsymbol{R}_\beta^0=\boldsymbol{R}_\beta\boldsymbol{T} \tag{36}$$

及由(35)式

$$\boldsymbol{R}_\beta^0=\begin{bmatrix}\boldsymbol{C}_1^0 & \boldsymbol{0}\\ \boldsymbol{C}_1^0\boldsymbol{A}_{11}^0 & \boldsymbol{0}\\ \vdots & \vdots\\ \boldsymbol{C}_1^0\boldsymbol{A}_{11}^{0\beta-1} & \boldsymbol{0}\end{bmatrix}\triangleq[\boldsymbol{R}_{\beta1}^0\,|\,\boldsymbol{0}],\ \operatorname{rank}\boldsymbol{R}_{\beta1}^0=q^0 \tag{37}$$

因此，可求出变换矩阵$\boldsymbol{T}$. 即，列变换

$$[\boldsymbol{R}_\beta]\longrightarrow[\boldsymbol{R}_{\beta1}^0\,|\,\boldsymbol{0}]$$

的变换矩阵给出$\boldsymbol{T}$. ”

现在假定，有两个系统S_1, S_2, 当任意指定S_1的起始状态时，与此对应就确定了S_2处于某个起始状态，当S_1, S_2处于这些起始状态时，对于任意输入两者的输出一致. 我们将此写成$S_1\subset S_2$. 当$S_1\subset S_2$且$S_2\subset S_1$时，因S_1和S_2的输入输出特性完全相同，则称两个系统有等价的输入输出关系，或简单地称为**等价**(equivalent)，表示成$S_1\sim S_2$.

因为定理6的$(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$和$(\boldsymbol{A}_{11}^0, \boldsymbol{B}_1^0, \boldsymbol{C}_1^0)$具有同样的传递函数，显然是零状态等价的，而且在上述意义上也是等价的. 这可说明如下.

两个系统的零输入响应

$$(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})\longrightarrow \boldsymbol{y}_{zi}=\boldsymbol{C}_e^{\boldsymbol{A}t}\boldsymbol{x}(0)$$

$$(\boldsymbol{A}_{11}^0, \boldsymbol{B}_1^0, \boldsymbol{C}_1^0)\longrightarrow \boldsymbol{y}_{zi}^0=\boldsymbol{C}_1^0e^{\boldsymbol{A}_{11}^0t}\boldsymbol{x}_1^0(0).$$

现在假定，在n维$(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$和$q^0$维$(\boldsymbol{A}_{11}^0, \boldsymbol{B}_1^0, \boldsymbol{C}_1^0)$中，起始状态$\boldsymbol{x}(0)$, $\boldsymbol{x}_1^0(0)$按下列关系给定，

$$\boldsymbol{x}_1^0(0)=\tilde{\boldsymbol{T}}_1\boldsymbol{x}(0),\quad 式中\ \boldsymbol{T}^{-1}\triangleq\begin{bmatrix}\tilde{\boldsymbol{T}}_1\\ \tilde{\boldsymbol{T}}_2\end{bmatrix} \tag{38}$$

这意味着，若给定$\boldsymbol{x}(0)$，则$\boldsymbol{x}_1^0(0)$由(38)式唯一确定，若给定$\boldsymbol{x}_1^0(0)$，则$\boldsymbol{x}(0)$不是唯一的，但一定要满足(38)式. 那么，若(38)式成立，则

$$\begin{aligned}\boldsymbol{y}_{zi}(t)&=\boldsymbol{C}e^{\boldsymbol{A}t}\boldsymbol{x}(0)=\boldsymbol{C}^0e^{\boldsymbol{A}^0t}\boldsymbol{T}^{-1}\boldsymbol{x}(0)\\&=\boldsymbol{C}_1^0e^{\boldsymbol{A}_{11}^0t}\boldsymbol{x}_1^0(0)=\boldsymbol{y}_{1zi}^0(t)\quad \forall t\geqslant 0\end{aligned}$$

因此，$(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$和$(\boldsymbol{A}_{11}^0, \boldsymbol{B}_1^0, \boldsymbol{C}_1^0)$关于零输入响应等价，而前面已经证明它们是零状态等价的. 因为输出可以用零输入响应和零状态响应之和给出，这就表明两个等价性都成立.

(问题4) 试证明等价性～是等值关系，即 $S_1 \sim S_1$; $S_1 \sim S_2$, $S_2 \sim S_3 \Rightarrow S_1 \sim S_3$; $S_1 \sim S_2 \Leftrightarrow S_2 \sim S_1$.

(问题5) 试证明在和 $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ 等价的标准形中定理6的 $(\boldsymbol{A}_{11}^0, \boldsymbol{B}_1^0, \boldsymbol{C}_1^0)$ 具有最小维数.

现在，将定理5和定理6结合起来，便可以得出卡尔曼典范结构定理，这是我们现在讨论的主要目的.

[定理7] 设 $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ 的可控性指标为 α，可观测性指标为 β，$\operatorname{rank} \boldsymbol{Q}_\alpha = p (\leqslant n)$，$\operatorname{rank} \boldsymbol{R}_\beta \boldsymbol{Q}_\alpha = q (\leqslant p)$. 这时，对于某个正则矩阵 $\boldsymbol{T} \in R^{n \times n}$，$(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ 和满足下列条件(i)，(ii)的 $(\boldsymbol{A}, \overline{\boldsymbol{B}}, \overline{\boldsymbol{C}})$ 代数等价.

(i)

$$\overline{\boldsymbol{A}} \triangleq \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \begin{array}{c} \begin{array}{cccc} \overbrace{\qquad}^{q} & \overbrace{\qquad}^{p-q} & \overbrace{\qquad\qquad\qquad}^{n-p} \end{array} \\ \left[\begin{array}{cc|cc} \overline{\boldsymbol{A}}_{11} & \boldsymbol{0} & \overline{\boldsymbol{A}}_{13} & \boldsymbol{0} \\ \overline{\boldsymbol{A}}_{21} & \overline{\boldsymbol{A}}_{22} & \boldsymbol{A}_{23} & \overline{\boldsymbol{A}}_{24} \\ \hline \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{A}_{33} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{A}_{43} & \boldsymbol{A}_{44} \end{array}\right] \begin{array}{l} \}q \\ \}p-q \\ \\ \}n-p \end{array} \end{array} \tag{39a}$$

$$\overline{\boldsymbol{B}} \triangleq \boldsymbol{T}^{-1}\boldsymbol{B} = \left[\begin{array}{c} \overline{\boldsymbol{B}}_1 \\ \overline{\boldsymbol{B}}_2 \\ \hline \boldsymbol{0} \\ \boldsymbol{0} \end{array}\right] \begin{array}{l} \}q \\ \}p-q \\ \\ \}n-q \end{array} \tag{39b}$$

$$\overline{\boldsymbol{C}} \triangleq \boldsymbol{C}\boldsymbol{T} = [\overbrace{\overline{\boldsymbol{C}}_1}^{q} \ \overbrace{\boldsymbol{0}}^{p-q} \mid \overbrace{\overline{\boldsymbol{C}}_3 \ \boldsymbol{0}}^{n-p}] \tag{39c}$$

(ii) 由子矩阵组成的 q 维 $(\boldsymbol{A}_{11}, \overline{\boldsymbol{B}}_1, \overline{\boldsymbol{C}}_1)$ 是可控且可观测的，和 $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ 具有同样的传递函数.

(证明) 因 $\operatorname{rank} \boldsymbol{Q}_\alpha = p$，由定理5，则对于某个 $\boldsymbol{T}_1 \in R^{n \times n}$，$(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ 和下列 $(\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{B}}, \tilde{\boldsymbol{C}})$ 代数等价:

$$\begin{aligned} \tilde{\boldsymbol{A}} &\triangleq \boldsymbol{T}_1^{-1}\boldsymbol{A}\boldsymbol{T}_1 = \begin{bmatrix} \tilde{\boldsymbol{A}}_{11} & \tilde{\boldsymbol{A}}_{12} \\ \boldsymbol{0} & \tilde{\boldsymbol{A}}_{22} \end{bmatrix} \begin{array}{l} \}p \\ \}n-p, \end{array} \\ \tilde{\boldsymbol{B}} &\triangleq \boldsymbol{T}_1^{-1}\boldsymbol{B} = \begin{bmatrix} \tilde{\boldsymbol{B}}_1 \\ \boldsymbol{0} \end{bmatrix} \begin{array}{l} \}p \\ \}n-p \end{array} \\ \tilde{\boldsymbol{C}} &\triangleq \boldsymbol{C}\boldsymbol{T}_1 = [\overbrace{\tilde{\boldsymbol{C}}_1}^{p} \ \overbrace{\tilde{\boldsymbol{C}}_2}^{n-p}] \end{aligned} \tag{40}$$

式中 $(\tilde{\boldsymbol{A}}_{11}, \tilde{\boldsymbol{B}}_1, \tilde{\boldsymbol{C}}_1)$ 是可控的，且和 $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ 零状态等价.

下面，从可观测性考虑，将 $(\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{B}}, \tilde{\boldsymbol{C}})$ 进行分解. 设 $(\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{B}}, \tilde{\boldsymbol{C}})$ 的可观测性矩阵为 $\tilde{\boldsymbol{R}}_\beta$，则由(40)式得

$$\tilde{\boldsymbol{R}}_\beta = \left[\begin{array}{c|c} \overbrace{\tilde{\boldsymbol{C}}_1}^{p} & \overbrace{\tilde{\boldsymbol{C}}_2}^{n-p} \\ \tilde{\boldsymbol{C}}_1\tilde{\boldsymbol{A}}_{11} & \tilde{\boldsymbol{C}}_1\tilde{\boldsymbol{A}}_{12} + \tilde{\boldsymbol{C}}_2\tilde{\boldsymbol{A}}_{22} \\ \vdots & \vdots \\ \tilde{\boldsymbol{C}}_1\tilde{\boldsymbol{A}}_{11}^{n-1} & \end{array}\right] \triangleq [\overbrace{\tilde{\boldsymbol{R}}_{\beta 1}}^{p} \mid \overbrace{\tilde{\boldsymbol{R}}_{\beta 2}}^{n-p}] \tag{41}$$

显然，$\widetilde{\boldsymbol{R}}_{\beta 1}$ 是 $(\widetilde{\boldsymbol{A}}_{11},\ \widetilde{\boldsymbol{B}}_1,\ \widetilde{\boldsymbol{C}}_1)$ 的可观测性矩阵．在这里若假定 $\operatorname{rank}\widetilde{\boldsymbol{R}}_\beta=\tilde{q}$，则在 $\widetilde{\boldsymbol{R}}_\beta$ 的 n 个列中有 $\tilde{q}$ 个线性独立组，设这些列是从第 1 列向右依次选取的．因 $\widetilde{\boldsymbol{R}}_\beta$ 中余下的 $(n-\tilde{q})$ 个列可以用上述 $\tilde{q}$ 列线性表出（B-I-3 中定理 2），则由适当的列变换矩阵

$$\begin{array}{cl} \begin{array}{cc} \scriptstyle p & \scriptstyle n-p \end{array} & \\ \boldsymbol{T}_2=\left[\begin{array}{c|c} \boldsymbol{T}_3 & \boldsymbol{T}_4 \\ \hline \boldsymbol{0} & \boldsymbol{T}_5 \end{array}\right] & \begin{array}{l} \}p \\ \}n-p \end{array} \end{array} \tag{42}$$

得到

$$\widetilde{\boldsymbol{R}}_\beta\boldsymbol{T}_2=[\overbrace{\widetilde{\boldsymbol{R}}_{\beta 1}\boldsymbol{T}_3}^{p}\mid\overbrace{\widetilde{\boldsymbol{R}}_{\beta 1}\boldsymbol{T}_4+\widetilde{\boldsymbol{R}}_{\beta 2}\boldsymbol{T}_5}^{n-p}]=[\overbrace{\overline{\boldsymbol{R}}_{\beta 1}}^{\hat{q}}\ \overbrace{\boldsymbol{0}}^{p-\hat{q}}\mid\overbrace{\boldsymbol{R}_{\beta 3}}^{\tilde{q}-\hat{q}}\ \boldsymbol{0}] \tag{43}$$

式中 $\hat{q}\triangleq\operatorname{rank}\widetilde{\boldsymbol{R}}_{\beta 1}$.

若由上面 $\boldsymbol{T}_2$ 变换 $(\widetilde{\boldsymbol{A}},\ \widetilde{\boldsymbol{B}},\ \widetilde{\boldsymbol{C}})$，则可以得到具有规定的(39)式形式的 $(\overline{\boldsymbol{A}},\ \overline{\boldsymbol{B}},\ \overline{\boldsymbol{C}})$．这一点可按如下理解．设将 $\widetilde{\boldsymbol{R}}_\beta$ 列变换成

$$\widetilde{\boldsymbol{R}}_\beta\widetilde{\boldsymbol{T}}_2=[\overbrace{\overline{\boldsymbol{R}}_{\beta 1}}^{\hat{q}},\overbrace{\overline{\boldsymbol{R}}_{\beta 3}}^{\tilde{q}-\hat{q}}\ \overbrace{\boldsymbol{0}}^{p-\hat{q}}\ \boldsymbol{0}] \tag{44}$$

的列变换矩阵为 $\widetilde{\boldsymbol{T}}_2$，根据定理 6，$\widetilde{\boldsymbol{A}}$，$\widetilde{\boldsymbol{C}}$ 可由 $\widetilde{\boldsymbol{T}}_2$ 变换成

$$\begin{array}{cl} \begin{array}{cccc} \scriptstyle \hat{q} & \scriptstyle \tilde{q}-\hat{q} & \scriptstyle p-\hat{q} & \end{array} & \\ \widetilde{\boldsymbol{T}}_2^{-1}\widetilde{\boldsymbol{A}}\widetilde{\boldsymbol{T}}_2=\left[\begin{array}{c|c|c|c} \hat{\boldsymbol{A}}_{11} & \hat{\boldsymbol{A}}_{12} & \boldsymbol{0} & \boldsymbol{0} \\ \hline \hat{\boldsymbol{A}}_{21} & \hat{\boldsymbol{A}}_{22} & \boldsymbol{0} & \boldsymbol{0} \\ \hline \hat{\boldsymbol{A}}_{31} & \hat{\boldsymbol{A}}_{32} & \hat{\boldsymbol{A}}_{33} & \hat{\boldsymbol{A}}_{34} \\ \hline \hat{\boldsymbol{A}}_{41} & \hat{\boldsymbol{A}}_{42} & \hat{\boldsymbol{A}}_{43} & \hat{\boldsymbol{A}}_{44} \end{array}\right] & \begin{array}{l} \}\tilde{q} \\ \}\tilde{q}-\hat{q} \\ \}p-\tilde{q} \\ \ \end{array} \end{array} \tag{45a}$$

$$\widetilde{\boldsymbol{C}}\boldsymbol{T}_2=[\overbrace{\hat{\boldsymbol{C}}_1}^{\hat{q}}\ \overbrace{\hat{\boldsymbol{C}}_2}^{\tilde{q}-\hat{q}}\ \overbrace{\boldsymbol{0}}^{p-\hat{q}}\ \boldsymbol{0}] \tag{45b}$$

但是，将(43)，(44)式比较后可以看到，$\boldsymbol{T}_2$ 和 $\widetilde{\boldsymbol{T}}_2$ 作用于同一个 $\widetilde{\boldsymbol{R}}_\beta$，得到结果之差别仅仅是第 2 个子块和第 3 个子块互相调换．即，$\boldsymbol{T}_2$ 是将 $\widetilde{\boldsymbol{T}}_2$ 的列作如下变换成的．

$$\boldsymbol{T}_2=\widetilde{\boldsymbol{T}}_2\begin{bmatrix} \boldsymbol{I}_{\hat{q}} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}_{\tilde{q}-\hat{q}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{\tilde{q}-\hat{q}} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}_\alpha \end{bmatrix}\quad(\alpha\triangleq n-p-\tilde{q}-\hat{q})$$

该式代入(45)式，则下面结果成立．

$$\overline{\boldsymbol{A}}\triangleq\boldsymbol{T}_2^{-1}\widetilde{\boldsymbol{A}}\boldsymbol{T}_2=\begin{bmatrix} \hat{\boldsymbol{A}}_{11} & \boldsymbol{0} & \hat{\boldsymbol{A}}_{12} & \boldsymbol{0} \\ \hat{\boldsymbol{A}}_{31} & \hat{\boldsymbol{A}}_{33} & \hat{\boldsymbol{A}}_{32} & \hat{\boldsymbol{A}}_{34} \\ \hat{\boldsymbol{A}}_{21} & \boldsymbol{0} & \hat{\boldsymbol{A}}_{22} & \boldsymbol{0} \\ \hat{\boldsymbol{A}}_{41} & \hat{\boldsymbol{A}}_{43} & \hat{\boldsymbol{A}}_{42} & \hat{\boldsymbol{A}}_{44} \end{bmatrix} \tag{46a}$$

$$\overline{\boldsymbol{C}}\triangleq\boldsymbol{C}\boldsymbol{T}_2=[\hat{\boldsymbol{C}}_1\ \ \boldsymbol{0}\ \ \hat{\boldsymbol{C}}_2\ \ \boldsymbol{0}] \tag{46b}$$

由(42)，(40)式可见，因 $\boldsymbol{T}_2$，$\widetilde{\boldsymbol{A}}$ 均为分块上三角形矩阵，则在(46a)式中显然 $\hat{\boldsymbol{A}}_{21}=\boldsymbol{0}$，$\hat{\boldsymbol{A}}_{41}=\boldsymbol{0}$，$\hat{\boldsymbol{A}}_{43}=\boldsymbol{0}$．而且，对于 $\boldsymbol{B}$ 直接由(40)，(42)式得

$$\overline{\boldsymbol{B}}\triangleq\boldsymbol{T}_2^{-1}\widetilde{\boldsymbol{B}}=\begin{bmatrix} \boldsymbol{T}_3^{-1}\widetilde{\boldsymbol{B}}_1 \\ \boldsymbol{0} \end{bmatrix}$$

在以上结果中，设

$$\begin{bmatrix} \hat{\boldsymbol{A}}_{11} & \mathbf{0} & \hat{\boldsymbol{A}}_{12} & \mathbf{0} \\ \hat{\boldsymbol{A}}_{31} & \hat{\boldsymbol{A}}_{33} & \hat{\boldsymbol{A}}_{32} & \hat{\boldsymbol{A}}_{34} \\ \mathbf{0} & \mathbf{0} & \hat{\boldsymbol{A}}_{22} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \hat{\boldsymbol{A}}_{42} & \hat{\boldsymbol{A}}_{44} \end{bmatrix} \triangleq \begin{bmatrix} \bar{\boldsymbol{A}}_{11} & \mathbf{0} & \bar{\boldsymbol{A}}_{13} & \mathbf{0} \\ \bar{\boldsymbol{A}}_{21} & \bar{\boldsymbol{A}}_{22} & \bar{\boldsymbol{A}}_{23} & \bar{\boldsymbol{A}}_{24} \\ \mathbf{0} & \mathbf{0} & \bar{\boldsymbol{A}}_{33} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \bar{\boldsymbol{A}}_{43} & \bar{\boldsymbol{A}}_{44} \end{bmatrix},$$

$$\left[\begin{array}{c} \boldsymbol{T}_3^{-1}\tilde{\boldsymbol{B}}_1 \\ \hline \mathbf{0} \end{array}\right] \triangleq \left[\begin{array}{c} \bar{\boldsymbol{B}}_1 \\ \hline \bar{\boldsymbol{B}}_2 \\ \hline \mathbf{0} \\ \mathbf{0} \end{array}\right], \quad [\tilde{\boldsymbol{C}}_1\ \mathbf{0}\ \tilde{\boldsymbol{C}}_2\ \mathbf{0}] \triangleq [\bar{\boldsymbol{C}}_1\ \mathbf{0}\ \bar{\boldsymbol{C}}_2\ \mathbf{0}],$$

便可得到(39)式形式．显然，可以利用 $\boldsymbol{T}\triangleq\boldsymbol{T}_1\boldsymbol{T}_2$ 作为从 $(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{C})$ 变换到最终的 $(\bar{\boldsymbol{A}},\ \bar{\boldsymbol{B}},\ \bar{\boldsymbol{C}})$ 的变换矩阵．而且，由证明过程及系数矩阵的形式可以清楚地看到，$(\bar{\boldsymbol{A}}_{11},\ \bar{\boldsymbol{B}}_1,\ \bar{\boldsymbol{C}}_1)$ 可控可观测，

$$\left(\begin{bmatrix} \bar{\boldsymbol{A}}_{11} & \mathbf{0} \\ \bar{\boldsymbol{A}}_{21} & \bar{\boldsymbol{A}}_{22} \end{bmatrix},\ \begin{bmatrix} \bar{\boldsymbol{B}}_1 \\ \bar{\boldsymbol{B}}_2 \end{bmatrix},\ [\bar{\boldsymbol{C}}_1\ \mathbf{0}]\right)\text{可控*},$$

$$\left(\begin{bmatrix} \bar{\boldsymbol{A}}_{22} & \bar{\boldsymbol{A}}_{24} \\ \mathbf{0} & \bar{\boldsymbol{A}}_{44} \end{bmatrix},\ \begin{bmatrix} \bar{\boldsymbol{B}}_2 \\ \mathbf{0} \end{bmatrix},\ [\mathbf{0}\ \mathbf{0}]\right)\text{不可观测},$$

$$\left(\begin{bmatrix} \bar{\boldsymbol{A}}_{11} & \bar{\boldsymbol{A}}_{12} \\ \mathbf{0} & \bar{\boldsymbol{A}}_{33} \end{bmatrix}\right),\ \begin{bmatrix} \bar{\boldsymbol{B}}_1 \\ \mathbf{0} \end{bmatrix},\ [\bar{\boldsymbol{C}}_1\ \boldsymbol{C}_2]\Bigg)\text{可观测},$$

$(\bar{\boldsymbol{A}}_{44},\ \mathbf{0},\ \mathbf{0})$ 不可控不可观测．

在这里，$\bar{\boldsymbol{A}}_{11}$ 的阶数 $\hat{q}$ 等于 $q\triangleq\text{rank}\,\boldsymbol{R}_\beta\boldsymbol{Q}_\alpha$ 必须证明，但这可考虑如下．由其构成方法看，$(\bar{\boldsymbol{A}}_{11},\ \bar{\boldsymbol{B}}_1,\ \bar{\boldsymbol{C}}_1)$ 是可控可观测的．设其可控性矩阵为 $\bar{\boldsymbol{Q}}_{\alpha 1}$，由[系 3]，$\text{rank}\,\bar{\boldsymbol{R}}_{\beta 1}\bar{\boldsymbol{Q}}_{\alpha 1}=\hat{q}$．但是，因 $(\bar{\boldsymbol{A}},\ \bar{\boldsymbol{B}},\ \bar{\boldsymbol{C}})$ 的可控性矩阵是

$$\bar{\boldsymbol{Q}}_\alpha=\begin{bmatrix} \bar{\boldsymbol{Q}}_{\alpha 1} \\ \bar{\boldsymbol{Q}}_{\alpha 2} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \begin{array}{l} \}\hat{q} \\ \}p-\tilde{q} \\ \left.\begin{array}{l} \\ \\ \end{array}\right\}n-p, \end{array}$$

可观测性矩阵 $\bar{\boldsymbol{R}}_\beta$ 可以用(43)式给出，则

$$\boldsymbol{R}_\beta\boldsymbol{Q}_\alpha=(\bar{\boldsymbol{R}}_\beta\boldsymbol{T}^{-1})(\boldsymbol{T}\bar{\boldsymbol{Q}}_\alpha)=\bar{\boldsymbol{R}}_\beta\bar{\boldsymbol{Q}}_\alpha=\bar{\boldsymbol{R}}_{\beta 1}\bar{\boldsymbol{Q}}_{\alpha 1}$$

因为由假定 $\text{rank}\,\boldsymbol{R}_\beta\boldsymbol{Q}_\alpha=q$，则必须 $\text{rank}\,\bar{\boldsymbol{R}}_{\beta 1}\bar{\boldsymbol{Q}}_{\alpha 1}=q$，即 $\hat{q}=q$．

最后来证明，$(\bar{\boldsymbol{A}}_{11},\ \bar{\boldsymbol{B}}_1,\ \bar{\boldsymbol{C}}_1)$ 和 $(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{C})$ 具有同样的传递函数(即零状态等价)．根据上面的证明过程，$(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{C})$ 和 $(\tilde{\boldsymbol{A}}_{11},\ \tilde{\boldsymbol{B}}_1,\ \tilde{\boldsymbol{C}})$ 是零状态等价的，而且 $(\tilde{\boldsymbol{A}}_{11},\ \tilde{\boldsymbol{B}}_1,\ \tilde{\boldsymbol{C}}_1)$ 和 $(\bar{\boldsymbol{A}}_{11},\ \bar{\boldsymbol{B}}_1,\ \bar{\boldsymbol{C}}_1)$ 零状态等价．因此，$(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{C})$ 和 $(\bar{\boldsymbol{A}}_{11},\ \bar{\boldsymbol{B}}_1,\ \bar{\boldsymbol{C}}_1)$ 也是零状态等价．

证明完毕

于是，该典范结构定理可解释如下．对应于(39)式的 $\bar{\boldsymbol{A}},\ \bar{\boldsymbol{B}},\ \bar{\boldsymbol{C}}$，将新的状态向量 $\bar{\boldsymbol{x}}(=\boldsymbol{T}^{-1}\boldsymbol{x})$ 分割写出如下，设

* 原文误印成 $\left(\begin{bmatrix} \boldsymbol{A}_{11} & \mathbf{0} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix},\ \begin{bmatrix} \bar{\boldsymbol{B}}_1 \\ \bar{\boldsymbol{B}}_2 \end{bmatrix},\ \bar{\boldsymbol{C}}[_1\mathbf{0}]\right)$ 可控．——译者注

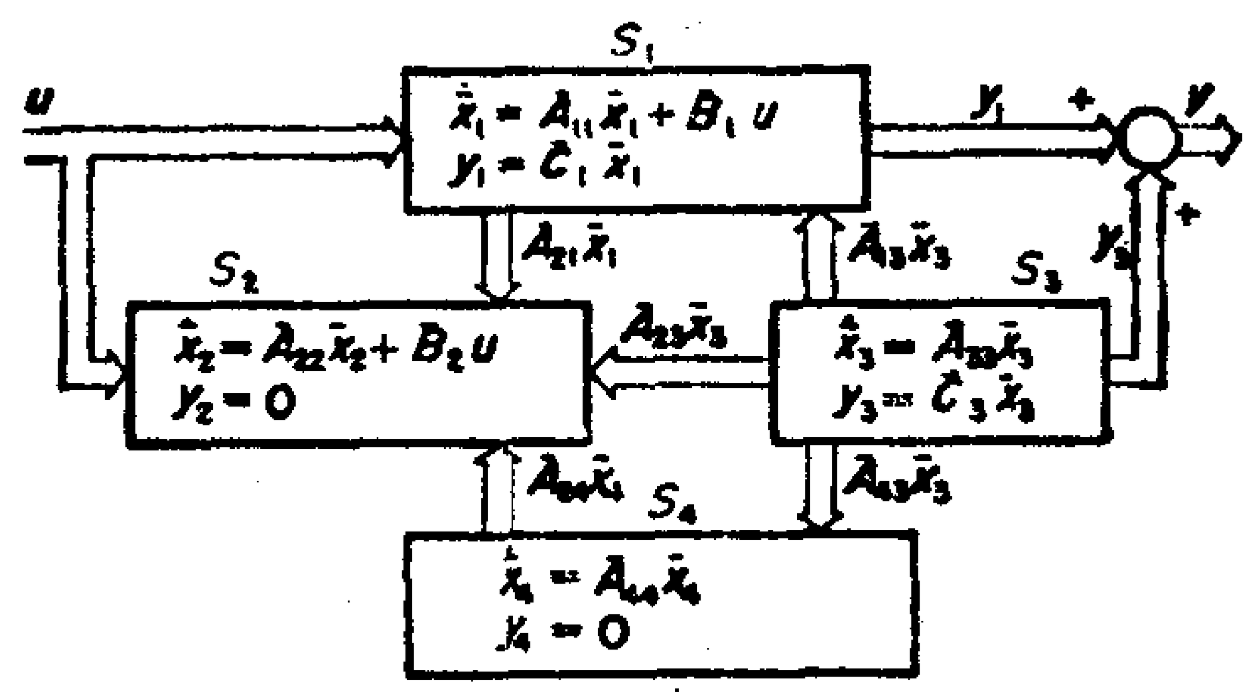


图 1　典范结构分解

$$\bar{\boldsymbol{x}}=\begin{bmatrix}\bar{\boldsymbol{x}}_1\\ \bar{\boldsymbol{x}}_2\\ \bar{\boldsymbol{x}}_3\\ \bar{\boldsymbol{x}}_4\end{bmatrix}\begin{matrix}\}q\\ \}p-q\\ \\ \}n-p\end{matrix}$$

则 $\dot{\bar{\boldsymbol{x}}}=\bar{\boldsymbol{A}}\bar{\boldsymbol{x}}+\bar{\boldsymbol{B}}\boldsymbol{u}$, $\boldsymbol{y}=\bar{\boldsymbol{C}}\bar{\boldsymbol{x}}$ 可用图 1 所示四个子系统 S_1, S_2, S_3, S_4 组合成的方框图表示．而且，关于各子系统的状态有如下性质：$\bar{\boldsymbol{x}}_1$＝可控"可观测"，$\bar{\boldsymbol{x}}_2$＝可控不可观测，$\bar{\boldsymbol{x}}_3$＝不可控"可观测"，$\bar{\boldsymbol{x}}_4$＝不可控不可观测．因零状态响应和 S_1 的输出 $\boldsymbol{y}_1$ 一致(注意，若 $\bar{\boldsymbol{x}}(0)=\boldsymbol{0}$，则 S_3, S_4 不激励，S_2 不连到输出*)，则传递函数可由 S_1 决定．

(注意)

"这里，所谓 $\bar{\boldsymbol{x}}_1$, $\bar{\boldsymbol{x}}_2$ 可控是指，

$$\bar{\boldsymbol{x}}(0)=\begin{bmatrix}\bar{\boldsymbol{x}}_1(0)\\ \bar{\boldsymbol{x}}_2(0)\\ \boldsymbol{0}\\ \boldsymbol{0}\end{bmatrix}$$

形的起始状态可控；$\bar{\boldsymbol{x}}_2$, $\bar{\boldsymbol{x}}_4$ 不可观测是指

$$\bar{\boldsymbol{x}}(0)=\begin{bmatrix}\boldsymbol{0}\\ \boldsymbol{x}_2(0)\\ \boldsymbol{0}\\ \boldsymbol{x}_4(0)\end{bmatrix}$$

形的起始状态不可观测(即产生恒等于零的零输入响应)．

而且，所谓 $\bar{\boldsymbol{x}}_1$, $\bar{\boldsymbol{x}}_3$ "可观测"是指，当预先知道起始状态为

$$\bar{\boldsymbol{x}}(0)=\begin{bmatrix}\bar{\boldsymbol{x}}_1(0)\\ \boldsymbol{0}\\ \bar{\boldsymbol{x}}_3(0)\\ \boldsymbol{0}\end{bmatrix}$$

形式时，$\bar{\boldsymbol{x}}(0)$ 完全由零输入响应决定．在给出一般的起始状态 $\bar{\boldsymbol{x}}(0)$ 的情况下，零输入响应仅能决定其 $\bar{\boldsymbol{x}}_1(0)$, $\bar{\boldsymbol{x}}_3(0)$ 分量，而 $\bar{\boldsymbol{x}}_2(0)$, $\bar{\boldsymbol{x}}_4(0)$ 是未知的．注意，象这样"可观测"意味着有条件的可观测．

根据该典范结构定理，当给定传递函数时，必须满足其最小实现的性质是显然的．

* 即若系统已处于零状态，那么输入 $\boldsymbol{u}$ 所能驱动的是子系统 S_1、S_2，S_1 的输出所能影响的只是 S_2，而 S_2 不能影响任何子系统．因此，S_3, S_4 仍处于零状态，S_2 不影响输出．——译者注

因此，再来讲几个直接由定理 7 得出的结果.

[系 6]　设给出真正适宜的以实系数有理函数为元素的 $m\times r$ 矩阵传递函数 $\boldsymbol{G}(s)$，则当且仅当 $(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{C})$ 可控可观测时，$\boldsymbol{G}(s)$ 的某个实现 $(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{C})$ 才是最小实现.

（证明）　设 $(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{C})$ 为可控可观测的实现. 若它不是最小实现，则应存在更小维数的实现 $(\tilde{\boldsymbol{A}},\ \tilde{\boldsymbol{B}},\ \tilde{\boldsymbol{C}})$. 因 $(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{C})$ 和 $(\tilde{\boldsymbol{A}},\ \tilde{\boldsymbol{B}},\ \tilde{\boldsymbol{C}})$ 零状态等价，则两者的传递函数（及脉冲响应）相同，则

$$\boldsymbol{C}\boldsymbol{A}^k\boldsymbol{B}=\tilde{\boldsymbol{C}}\tilde{\boldsymbol{A}}^k\tilde{\boldsymbol{B}},\ \ k=0,\ 1,\ 2,\ \cdots\cdots$$

因此，当设 $(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{C})$ 的可控性矩阵，可观测性矩阵分别为 $\boldsymbol{Q},\ \boldsymbol{R}$ 时，

$$\boldsymbol{R}\boldsymbol{Q}=\begin{bmatrix}\tilde{\boldsymbol{C}}\\ \tilde{\boldsymbol{C}}\tilde{\boldsymbol{A}}\\ \vdots\\ \tilde{\boldsymbol{C}}\tilde{\boldsymbol{A}}^{(n-1)}\end{bmatrix}[\tilde{\boldsymbol{B}}\ \ \tilde{\boldsymbol{A}}\tilde{\boldsymbol{B}}\cdots\tilde{\boldsymbol{A}}^{(n-1)}\tilde{\boldsymbol{B}}]$$

成立. 但是根据[系 3]，$\operatorname{rank}\boldsymbol{R}\boldsymbol{Q}=n$，这与上式矛盾. 因此，不存在维数更小的实现 $(\tilde{\boldsymbol{A}},\ \tilde{\boldsymbol{B}},\ \tilde{\boldsymbol{C}})$.

反之，设 $(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{C})$ 为最小实现. 若它不具有可控可观测的性质（即至少其中有一个不成立），由定理 7，则可以找出低维数且零状态等价的 $(\boldsymbol{A}_{11},\ \bar{\boldsymbol{B}}_1,\ \bar{\boldsymbol{C}}_1)$，这与最小实现的假定相矛盾.

证明完毕

[系 7]　给出真正适宜的以实系数有理函数为元素的 $m\times r$ 矩阵传递函数 $\boldsymbol{G}(s)$. 这时，$\boldsymbol{G}(s)$ 的最小实现都是代数等价的，即在任意两个最小实现 $(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{C})$ 和 $(\tilde{\boldsymbol{A}},\ \tilde{\boldsymbol{B}},\ \tilde{\boldsymbol{C}})$ 之间，对于某个正则矩阵 $\boldsymbol{T}$，有 $\tilde{\boldsymbol{A}}=\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T},\ \tilde{\boldsymbol{B}}=\boldsymbol{T}^{-1}\boldsymbol{B},\ \tilde{\boldsymbol{C}}=\boldsymbol{C}\boldsymbol{T}$ 的关系存在.

（证明）　由[系 6]，$(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{C})$ 及 $(\tilde{\boldsymbol{A}},\ \tilde{\boldsymbol{B}},\ \tilde{\boldsymbol{C}})$ 是可控可观测的. 因此，$\operatorname{rank}\boldsymbol{Q}=\operatorname{rank}\boldsymbol{R}=\operatorname{rank}\tilde{\boldsymbol{Q}}=\operatorname{rank}\tilde{\boldsymbol{R}}=n$. 而且由 $(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{C})$ 和 $(\tilde{\boldsymbol{A}},\ \tilde{\boldsymbol{B}},\ \tilde{\boldsymbol{C}})$ 的零状态等价性，则 $\boldsymbol{C}\boldsymbol{A}^k\boldsymbol{B}=\tilde{\boldsymbol{C}}\tilde{\boldsymbol{A}}^k\tilde{\boldsymbol{B}},\ k=0,\ 1,\ 2,\ \cdots$，故得 $\boldsymbol{R}\boldsymbol{Q}=\tilde{\boldsymbol{R}}\tilde{\boldsymbol{Q}}$. 因此，$\tilde{\boldsymbol{R}}^T\boldsymbol{R}\boldsymbol{Q}=\tilde{\boldsymbol{R}}^T\tilde{\boldsymbol{R}}\tilde{\boldsymbol{Q}}$ 成立.

因 $\tilde{\boldsymbol{R}}^T\boldsymbol{R}\in R^{n\times n}$ 是正则的（B-I-3, 问题 2），则

$$\tilde{\boldsymbol{Q}}=(\tilde{\boldsymbol{R}}^T\tilde{\boldsymbol{R}})^{-1}\tilde{\boldsymbol{R}}^T\boldsymbol{R}\boldsymbol{Q}\tag{47}$$

这里，因 $\operatorname{rank}\tilde{\boldsymbol{Q}}=\operatorname{rank}\boldsymbol{Q}=n$，则 $n\times n$ 矩阵 $(\tilde{\boldsymbol{R}}^T\tilde{\boldsymbol{R}})^{-1}\tilde{\boldsymbol{R}}^T\boldsymbol{R}$ 是正则的，因此，令 $\boldsymbol{T}_1^{-1}\triangleq(\tilde{\boldsymbol{R}}^T\boldsymbol{R})^{-1}\tilde{\boldsymbol{R}}^T\boldsymbol{R}$. 再由 $\boldsymbol{R}\boldsymbol{Q}=\tilde{\boldsymbol{R}}\tilde{\boldsymbol{Q}}$，和上面同样，可以得

$$\tilde{\boldsymbol{R}}=\boldsymbol{R}\boldsymbol{Q}\tilde{\boldsymbol{Q}}^T(\tilde{\boldsymbol{Q}}\tilde{\boldsymbol{Q}}^T)^{-1}\tag{48}$$

在这里令 $\boldsymbol{T}_2\triangleq\boldsymbol{Q}\tilde{\boldsymbol{Q}}^T(\tilde{\boldsymbol{Q}}\tilde{\boldsymbol{Q}}^T)^{-1}$，则 $\boldsymbol{T}_2$ 是正则的，考虑到 $\boldsymbol{R}\boldsymbol{Q}=\tilde{\boldsymbol{R}}\tilde{\boldsymbol{Q}}$，得

$$\boldsymbol{T}_1^{-1}\boldsymbol{T}_2=\{(\tilde{\boldsymbol{R}}^T\tilde{\boldsymbol{R}})^{-1}\tilde{\boldsymbol{R}}^T\tilde{\boldsymbol{R}}\}\{\boldsymbol{Q}\tilde{\boldsymbol{Q}}^T(\tilde{\boldsymbol{Q}}\tilde{\boldsymbol{Q}}^T)^{-1}\}=\boldsymbol{I}_n$$

即 $\boldsymbol{T}_1=\boldsymbol{T}_2$. 因此，重新写成 $\boldsymbol{T}\triangleq\boldsymbol{T}_1=\boldsymbol{T}_2$.

(47)式可以表示成 $\tilde{\boldsymbol{Q}}=\boldsymbol{T}^{-1}\boldsymbol{Q}$，比较 $\tilde{\boldsymbol{Q}},\ \boldsymbol{Q}$ 的第 1 个子块，得 $\tilde{\boldsymbol{B}}=\boldsymbol{T}^{-1}\boldsymbol{B}$. 同样，(48)式可以表示成 $\tilde{\boldsymbol{R}}=\boldsymbol{R}\boldsymbol{T}$，由此 $\tilde{\boldsymbol{C}}=\boldsymbol{C}\boldsymbol{T}$. 而且可以看到，因 $\boldsymbol{C}\boldsymbol{A}^k\boldsymbol{B}=\tilde{\boldsymbol{C}}\tilde{\boldsymbol{A}}^k\tilde{\boldsymbol{B}},\ k=0,\ 1,\ 2,\ \cdots$，则 $\boldsymbol{R}\boldsymbol{Q}=\tilde{\boldsymbol{R}}\tilde{\boldsymbol{Q}}$ 和

$$\boldsymbol{R}\boldsymbol{A}\boldsymbol{Q}=\tilde{\boldsymbol{R}}\tilde{\boldsymbol{A}}\tilde{\boldsymbol{Q}}\tag{49}$$

同时成立. 由此可以证明

$$\widetilde{\boldsymbol{A}}=(\widetilde{\boldsymbol{R}}^T\widetilde{\boldsymbol{R}})^{-1}\widetilde{\boldsymbol{R}}^T\boldsymbol{RAQ}\widetilde{\boldsymbol{Q}}^T(\widetilde{\boldsymbol{Q}}\widetilde{\boldsymbol{Q}}^T)^{-1}=\boldsymbol{T}^{-1}\boldsymbol{AT} \tag{50}$$

证明完毕

由上述系 6, 7 可见，最小实现的性质是显然的．亦即，在给出真正适宜的有理传递函数矩阵 $\boldsymbol{G}(s)$ 的情况下，则当且仅当 $(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{C})$ 可控可观测时，其实现之一 $(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{C})$ 才是最小实现．改变正则矩阵 $\boldsymbol{T}\in R^{n\times n}$（同时保持其正则性），可以用 $(\boldsymbol{T}^{-1}\boldsymbol{AT},\ \boldsymbol{T}^{-1}\boldsymbol{B},\ \boldsymbol{CT})$ 的形式，由 $(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{C})$ 产生 $\boldsymbol{G}(s)$ 的所有最小实现．为了求取其中的一个最小实现，例如，可以用 A-I-10 中的实现方法求出实现 $(\widetilde{\boldsymbol{A}},\ \widetilde{\boldsymbol{B}},\ \widetilde{\boldsymbol{C}})$（一般不是最小实现），再利用定理 7 分解成典型结构，在其中取出和 $(\widetilde{\boldsymbol{A}},\ \widetilde{\boldsymbol{B}},\ \widetilde{\boldsymbol{C}})$ 零状态等价的可控可观测部分 $(\boldsymbol{A}_{11},\ \overline{\boldsymbol{B}}_1,\ \overline{\boldsymbol{C}}_1)$ 即可．即，由 $(\widetilde{\boldsymbol{A}},\ \widetilde{\boldsymbol{B}},\ \widetilde{\boldsymbol{C}})$ 中去掉不可控不可观测部分，使状态空间的维数降到最小值．

在没有讲到 $\boldsymbol{G}(s)$ 的麦克米兰次数等之前，还不能详细地讨论求最小实现更直接的方法，至于代数方法可参阅文献[82]中的说明．

（注意）

以上在讲述典范结构定理及其有关内容时，是假定(30)式中的 $\boldsymbol{D}=\boldsymbol{0}$．即使 $\boldsymbol{D}\neq\boldsymbol{0}$，也不会有本质性影响，仅仅是在图 1 所示方框图中输入输出之间插入直接通路．既使对于实现问题，若 $\boldsymbol{G}(s)$ 是适宜的，令 $\lim\limits_{s\to\infty}\boldsymbol{G}(s)\triangleq\boldsymbol{D}$，则 $\widetilde{\boldsymbol{G}}(s)\triangleq\boldsymbol{G}(s)-\boldsymbol{D}$ 亦是真正适宜的．因此，设 $\widetilde{\boldsymbol{G}}(s)$ 的一个（最小）实现为 $(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{C})$，则 $\boldsymbol{G}(s)$ 的（最小）实现是 $(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{C},\ \boldsymbol{D})$．即，$\boldsymbol{G}(s)$ 不能唯一确定 $\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{C}$，而 $\boldsymbol{D}$ 是唯一的．

（问题 6） 试证明具有纯量输入输出的 $(\boldsymbol{A},\ \boldsymbol{b},\ \boldsymbol{c},\ d)$ 的传递函数

$$g(s)=\boldsymbol{c}[s\boldsymbol{I}-\boldsymbol{A}]^{-1}\boldsymbol{b}+d=\frac{\boldsymbol{c}\,\mathrm{adj}[s\boldsymbol{I}-\boldsymbol{A}]\boldsymbol{b}}{\det[s\boldsymbol{I}-\boldsymbol{A}]}+d$$

的分子分母多项式，当且仅当 $(\boldsymbol{A},\ \boldsymbol{b},\ \boldsymbol{c},\ d)$ 可控且可观测时，不具有公共因子．

输入输出关系的反演及可控性、可观测性

在 A-I-4 中已经讨论过其特性用(30)式标准形表示的子系统组成的方框图．根据在该方框图中将输入输出关系反演得到的系统和反演前的系统之间关系的研究可以知道[83]，系统的分析、设计可以统一讨论．因此，最后再稍讲一下这个关系．在这里，输入输出关系的反演是指如下两个操作（参看图 2）．

(i) 信号流动的方向反转，即将方框图中箭头的方向反转．其结果是，各方框的输入变为输出，输出变为输入．同样，整个系统的输入变为输出，输出变为输入．而且，相加点

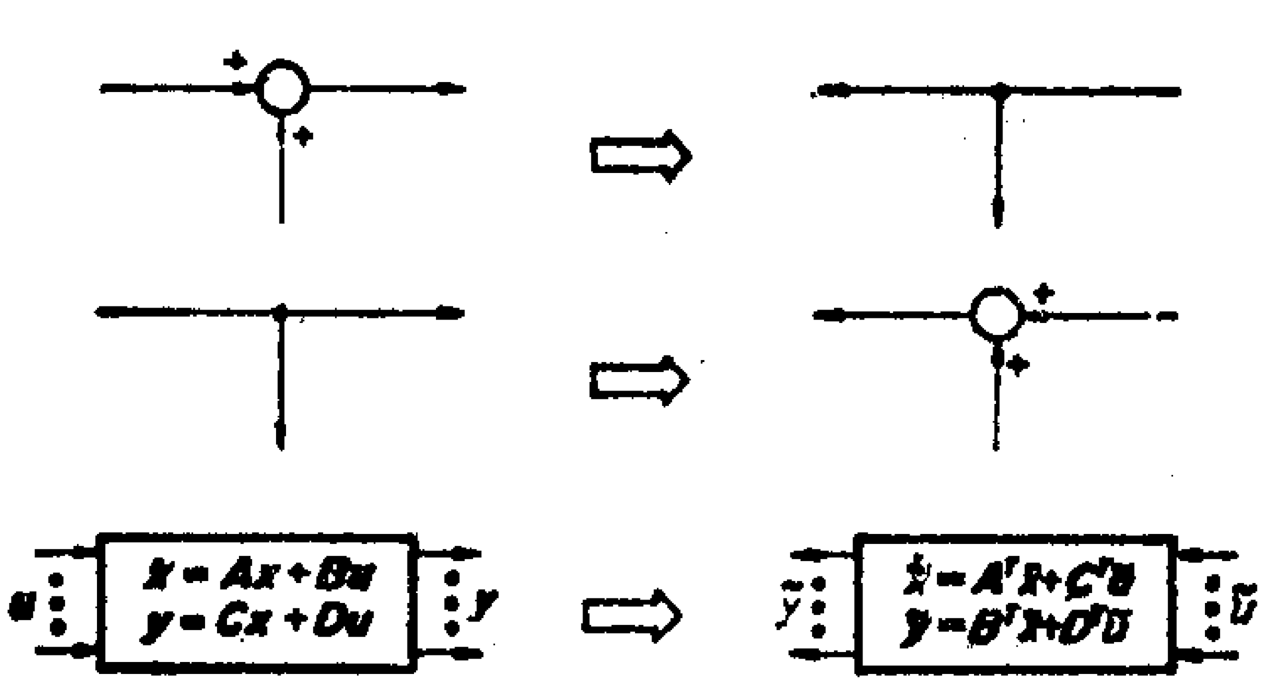


图 2 输入输出关系反演操作

变为引出点，引出点变成相加点[1]．

(ii) 将各方框用其对偶系统置换．即，将

$$\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{u},\ \ \boldsymbol{y}=\boldsymbol{C}\boldsymbol{x}+\boldsymbol{D}\boldsymbol{u} \tag{30}$$

换成对偶的标准形

$$\dot{\tilde{\boldsymbol{x}}}=\boldsymbol{A}^T\tilde{\boldsymbol{x}}+\boldsymbol{C}^T\tilde{\boldsymbol{u}},\ \ \tilde{\boldsymbol{y}}=\boldsymbol{B}^T\tilde{\boldsymbol{x}}+\boldsymbol{D}^T\tilde{\boldsymbol{u}}. \tag{51}$$

在这里要注意，根据(i)的操作，$\tilde{\boldsymbol{u}}$ 是 $\boldsymbol{y}$ 置换的，$\tilde{\boldsymbol{y}}$ 是 $\boldsymbol{u}$ 置换的．

这时要注意，若各方框的编号、输入输出编号都不改变[2]，则新系统的方框-方框，输入-方框，方框-输出，输入-输出的关联矩阵 $\tilde{\boldsymbol{F}}$, $\tilde{\boldsymbol{G}}$, $\tilde{\boldsymbol{J}}$, $\tilde{\boldsymbol{K}}$ 和原系统的关联矩阵 $\boldsymbol{F}$, $\boldsymbol{G}$, $\boldsymbol{J}$, $\boldsymbol{K}$ 之间有下列关系

$$\begin{aligned}\tilde{\boldsymbol{F}}&=\boldsymbol{F}^T,\ \ \tilde{\boldsymbol{G}}=\boldsymbol{J}^T\\ \tilde{\boldsymbol{J}}&=\boldsymbol{G}^T,\ \ \tilde{\boldsymbol{K}}=\boldsymbol{K}^T.\end{aligned} \tag{52}$$

这些关系可以证明如下．

现在，我们假定 $\boldsymbol{F}$ 的 (ij) 子块的 (kl) 元素为 1(0)[3]．这意味着原系统中 S_j 的第 l 个输出连接(不连接) S_i 的第 k 个输入．因此，在将信号流动方向反转后得到的新系统中，S_i 的第 k 个输出连接(不连接)S_j 的第 l 个输入，由 $\tilde{\boldsymbol{F}}$ 的定义，其 (ji) 子块的 (lk) 元素是 1(0)，则 $\tilde{\boldsymbol{F}}=\boldsymbol{F}^T$．其它关系也可以同样证明．

利用这些关系可以看到，在新系统中

$$\tilde{\boldsymbol{U}}=\boldsymbol{F}^T\tilde{\boldsymbol{Y}}+\boldsymbol{J}^T\tilde{\boldsymbol{V}} \tag{53}$$

$$\tilde{\boldsymbol{Z}}=\boldsymbol{G}^T\tilde{\boldsymbol{Y}}+\boldsymbol{K}^T\tilde{\boldsymbol{V}} \tag{54}$$

成立．式中 $\tilde{\boldsymbol{U}}$, $\tilde{\boldsymbol{Y}}$ 分别为各子块的新输入 $\tilde{\boldsymbol{u}}_i$，输出 $\tilde{\boldsymbol{y}}_i$ $(i=1,\ 2,\ \cdots,\ p)$ 纵向排成的实列向量．$\tilde{\boldsymbol{U}}$, $\tilde{\boldsymbol{Z}}$ 是整个系统新的输入及输出．这些 $\tilde{\boldsymbol{U}}$, $\tilde{\boldsymbol{Y}}$, $\tilde{\boldsymbol{V}}$. $\tilde{\boldsymbol{Z}}$ 是将原系统的 $\boldsymbol{Y}$, $\boldsymbol{U}$, $\boldsymbol{Z}$, $\boldsymbol{V}$ 置换成的．

另一方面，因方框新的输入输出特性是原方框的对偶系统，若不改变方框的编号，则 $S_i(i=1,\ 2,\ \cdots,\ p)$ 的特性为

$$\begin{aligned}\dot{\tilde{\boldsymbol{X}}}_i&=\boldsymbol{A}_i^T\tilde{\boldsymbol{x}}_i+\boldsymbol{C}_i^T\tilde{\boldsymbol{u}}_i\\ \tilde{\boldsymbol{Y}}_i&=\boldsymbol{B}_i^T\tilde{\boldsymbol{x}}_i+\boldsymbol{D}_i^T\tilde{\boldsymbol{u}}_i\end{aligned}$$

式中 $\tilde{\boldsymbol{x}}_i$ 是 S_i 新的状态．设 $\tilde{\boldsymbol{X}}$ 是 $\tilde{\boldsymbol{x}}_i(i=1,\ 2,\ \cdots,\ p)$ 纵向排成的实列向量，利用 $\tilde{\boldsymbol{X}}$ 等将上式归纳后可写成

$$\dot{\tilde{\boldsymbol{X}}}=\boldsymbol{A}^T\tilde{\boldsymbol{X}}+\boldsymbol{C}^T\tilde{\boldsymbol{U}} \tag{55a}$$

$$\tilde{\boldsymbol{Y}}=\boldsymbol{B}^T\tilde{\boldsymbol{X}}+\boldsymbol{D}^T\tilde{\boldsymbol{U}}. \tag{55b}$$

由以上可见，将信号流动方向反转后得到的新系统，可以用原系统中定义的矩阵表示．并且可以看到，在表示该新系统的(53)，(54)，(55)式中若应用 A-I-4 中定理 3，则 $\det[\boldsymbol{I}-\boldsymbol{D}^T\boldsymbol{F}^T]\ (=\det[\boldsymbol{I}-\boldsymbol{F}^T\boldsymbol{D}^T])\neq 0$ 是以 $\tilde{\boldsymbol{X}}$ 为状态整个系统可以用标准形完全表示的充分必要条件．但是，如 P-I-2 中所述，因为矩阵转置后其行列式值不变，则条件 $\det[\boldsymbol{I}$

1) 为了使相加点换成引出点时不产生矛盾，原系统相加点输入的符号必须全为正，否则要前置以放大倍数为 (-1) 的放大器．

2) 例如，原系统中 S_i 的第 k 个输入(输出)变成新系统中 S_i 的第 k 个输出(输入)．对于整个系统的输入和输出也是同样．

3) 如以前注解中所述，这里只考虑以 0 或 1 为元素的 $\boldsymbol{F}$, $\boldsymbol{G}$, $\boldsymbol{J}$, $\boldsymbol{K}$．

$-D^TF^T]\neq 0$ 和条件 $\det[I-FD]\neq 0$ 一样．后者是原整个系统可以用以 X 为状态的标准形完全表示的充分必要条件．因此，当新的整个系统可以用以 $\tilde{X}$ 为状态的标准形完全表示时，则原整个系统也可以用以 X 为状态的标准形完全表示，其逆亦成立．这时，将 A-I-4中(5)′，(6)′，(8)′式和(53)，(54)，(55)式进行比较，便可由 A-I-4 中(11)式求出表示新的整个系统的标准形状态方程式和输出方程式如下

$$\dot{\tilde{X}}=\{A^T+C^T(I-F^TD^T)^{-1}F^TB^T\}\tilde{X}+C^T(I-F^TD^T)^{-1}J^T\tilde{V} \tag{56a}$$

$$\tilde{Z}=G^T(I-D^TF^T)^{-1}B^T\tilde{X}+\{K^T+G^T(I-D^TF^T)D^TJ^T\}\tilde{V} \tag{56b}$$

(问题 7)　试证明(56)式系统是 A-I-4 中(11)式的对偶系统，它们的传递函数矩阵互为转置关系．

提示：
$$(I-FD)^{-1}F=F(I-DF)^{-1},$$
$$(I-DF)^{-1}D=D(I-FD)^{-1}.$$

将以上归纳，得如下定理．

［定理 8］　将输入输出关系反演后得到的整个系统，可以以子块状态的直和为状态，用标准形状态方程式和输出方程式完全表示的充分必要条件是，反演前的原系统也可以用同样的方式表示．当可能用这样的标准形表示时，将输入输出关系反演后得到的整个系统，和反演前的整个系统互为对偶系统．反演后系统的传递函数矩阵是反演前系统的传递函数矩阵的转置．因此，若反演前的系统可控，则反演后的系统可观测；若反演前的系统可观测，则反演后的系统可控．

这样，将输入输出关系反演后得到的系统是原系统的对偶系统，其传递函数矩阵是原系统传递函数矩阵的转置．因此，在用方框图表示的系统中设置补偿器，设计稳定系统等情况下，可以按照容易处理的信号流动方向进行设计．关于这一点详见文献[83]．

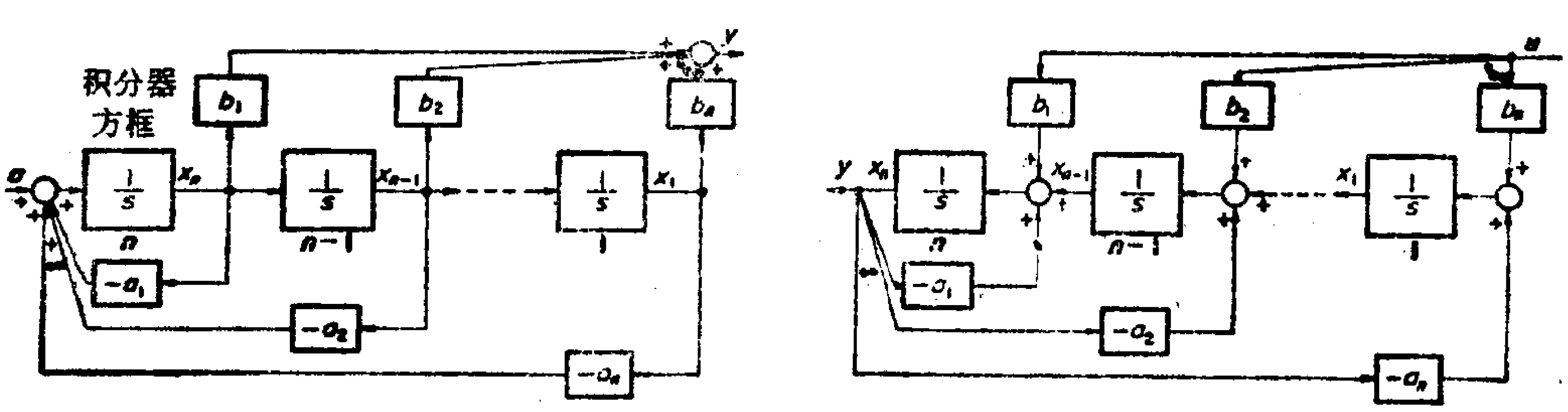
(例 5)[28]　在用模拟计算机实现给定传递函数的情况下，常使用这样两种实现方法，一种是用所谓的可控标准形，另一种是用所谓的可观测标准形的形式实现．现在假定传递函数用

$$\frac{b_1s^{n-1}+\cdots+b_n}{s^n+a_1s^{n-1}+\cdots+a_n} \tag{57}$$

给出，图 3(a)是用可控标准形实现的方框图，图 3(b)是用可观测标准形实现的方框图．这里要注意，若设状态为 x，输入为 u，输出为 y，用标准形

$$\dot{x}=u,\quad y=x$$

表示积分器的输入输出特性，则其本身就互为对偶系统．象这样地表示积分器的特性，显



(a)用可控标准形实现　　　(b)用可观测标准形实现

图 3　传递函数 $\dfrac{b_1s^{n-1}+\cdots+b_n}{s^n+a_1s^{n-1}+\cdots+a_n}$ 的实现

然，将图 3(a) 方框图输入输出关系反演后便是图 3(b) 的方框图(因为这两个系统是单输入单输出系统，故其传递函数一致)．在该二图中，若设第 i 个积分器方框的状态为 x_i，因 (a) 系统可以用标准形

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_{n-1} \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ -a_n & -a_{n-1} & \cdots & -a_2 & -a_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} u \tag{58a}$$

$$y = [\, b_n \ b_{n-1} \ \cdots \ b_2 b_1 \,] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} \tag{58b}$$

表示，(b) 系统可用标准形

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_{n-1} \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_n \\ 1 & 0 & \cdots & 0 & -a_{n-1} \\ \vdots & \ddots & & \vdots & \vdots \\ 0 & 0 & \ddots & 0 & -a_2 \\ 0 & 0 & \cdots & 1 & -a_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} b_n \\ b_{n-1} \\ \vdots \\ b_2 \\ b_1 \end{bmatrix} u \tag{59a}$$

$$y = [\, 0 \ 0 \ \cdots \ 0 \ 1 \,] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} \tag{59b}$$

表示，则 (b) 系统确实是 (a) 系统的对偶系统．显而易见，(a) 系统可控，(b) 系统可观测，根据它们的表示式 (58)，(59) 系数矩阵的结构，(a) 称为用可控标准形实现，(b) 称为用可观测标准形实现．因当且仅当该二系统传递函数 (57) 式分子分母不具有公共因子时，(a)，(b) 系统是可控且可观测的(参看问题 6)，则若具有公共因子，可控的 (a) 系统不可观测，可观测的 (b) 系统不可控．

(问题 8) 试证明图 3(a) 系统可用 (58) 式标准形表示，图 3(b) 系统可用 (59) 式标准形表示．

(问题 9) 试证明用标准形

$$\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{u}, \quad \boldsymbol{y} = \boldsymbol{C}\boldsymbol{x} + \boldsymbol{D}\boldsymbol{u}$$

表示的系统的传递函数矩阵，是其对偶系统

$$\dot{\boldsymbol{x}} = \boldsymbol{A}^T\boldsymbol{x} + \boldsymbol{C}^T\boldsymbol{u}, \quad \boldsymbol{y} = \boldsymbol{B}^T\boldsymbol{x} + \boldsymbol{D}^T\boldsymbol{u}$$

的传递函数矩阵的转置．

附录

A-I-8 中引理 1 (第 212 页) 证明不充分之处，这里再补充一下．

1．因为所有的矩阵范数$\|\boldsymbol{A}\|$都相互等价(B-I-5 中问题 7)，所以对于某个任意的范数$\|\cdot\|$，若有满足$\|\boldsymbol{A}(t)\|<k'$，$\|\boldsymbol{B}(t)\|<k'$，$\forall t\in R$ 的常数 k' 存在，则对于欧几里得范数$\|\cdot\|_2$，也有满足$\|\boldsymbol{A}(t)\|_2<k$，$\|\boldsymbol{B}(t)\|_2<k$，$\forall t\in R$ 的常数 k 存在．基于该结果，我们在引理 1 的证明中将使用欧几里得范数．

对于欧几里得范数，设 $\boldsymbol{S}$ 为任意 $m\times n$ 矩阵，利用 B-I-5 中(13)式及 B-I-11 中系 10，可以证明

$$\|\boldsymbol{S}\|_2=\max_i[\lambda_i(\boldsymbol{S}^T\boldsymbol{S})]^{1/2}=\max_i[\lambda_i(\boldsymbol{S}\boldsymbol{S}^T)]^{1/2}=\|\boldsymbol{S}^T\|_2.$$

该性质经常使用(例如，推导(40)式)．

2．(43)式可推导如下．由(39)式及 B-I-11 中系 22，得

$$\boldsymbol{W}_c^{-1}(t,\ t+\sigma)\leqslant\frac{1}{\alpha_0(\sigma)}\boldsymbol{I}\tag{1}$$

因$[\boldsymbol{\Phi}(t+\sigma,\ t)]^{-1}=\boldsymbol{\Phi}(t,\ t+\sigma)$是正则的，则由(1)式得

$$\boldsymbol{\Phi}(t,\ t+\sigma)^T\boldsymbol{W}_c^{-1}(t,\ t+\sigma)\boldsymbol{\Phi}(t,\ t+\sigma)\leqslant\frac{1}{\alpha_0(\sigma)}\boldsymbol{\Phi}(t,\ t+\sigma)^T\boldsymbol{\Phi}(t,\ t+\sigma).\tag{2}$$

其中因

$$\|\boldsymbol{\Phi}(t,\ t+\sigma)\|\leqslant\exp k\sigma$$

成立(注意(i))，则对于 $\beta_0(\sigma)\triangleq\alpha_0(\sigma)\exp(-2k\sigma)$，有

$$\frac{1}{\alpha_0(\sigma)}\|\boldsymbol{\Phi}(t,\ t+\sigma)^T\boldsymbol{\Phi}(t,\ t+\sigma)\|\leqslant\frac{1}{\beta_0(\sigma)}.$$

利用注意(ii)的性质，得

$$\frac{1}{\alpha_0(\sigma)}\boldsymbol{\Phi}(t,\ t+\sigma)^T\boldsymbol{\Phi}(t,\ t+\sigma)\leqslant\frac{1}{\beta_0(\sigma)}\boldsymbol{I}\tag{3}$$

由(2)，(3)式，得

$$\frac{1}{\beta_0(\sigma)}\boldsymbol{I}\geqslant\boldsymbol{\Phi}(t,\ t+\sigma)^T\boldsymbol{W}_c^{-1}(t,\ t+\sigma)\boldsymbol{\Phi}(t,\ t+\sigma)$$

对上式使用 B-I-11 中系 22 便得到(43)式．

3．注意：不限于欧几里得范数，对于一般的由矩阵导出的矩阵范数也有下列性质．

(i) 若$\|\boldsymbol{A}(t)\|<k$，则$\|\boldsymbol{\Phi}(t,\ s)\|\leqslant\exp k|t-s|$，$\forall t,\ s\in R$．其中 $\boldsymbol{\Phi}(t,\ s)$是 $\boldsymbol{A}(t)$ 的状态转移矩阵．

(证明) 对于 $t\geqslant s$ 的情况，A-I-6 中(22)式已证过，这里只考虑 $t\leqslant s$ 的情况．

将

$$\frac{\partial}{\partial t}\boldsymbol{\Phi}(t,\ s)=\boldsymbol{A}(t)\boldsymbol{\Phi}(t,\ s)$$

两边积分，得

$$\boldsymbol{\Phi}(t,\ s)=\boldsymbol{I}-\int_t^s\boldsymbol{A}(\tau)\boldsymbol{\Phi}(\tau,\ s)d\tau$$

因此，考虑到$\|\boldsymbol{I}\|=1$，则

$$\|\boldsymbol{\Phi}(t,\ s)\|\leqslant1+k\int_t^s\|\boldsymbol{\Phi}(\tau,\ s)\|d\tau=u(t,\ s)\ (\forall t\leqslant s)\tag{4}$$

这里若令(4)式右边为 $u(t,\ s)$，则

$$\frac{\partial}{\partial t} u(t, s)e^{k(t-s)} = ke^{k(t-s)}\{u(t, s) - \|\boldsymbol{\Phi}(t, s)\|\} \geqslant 0, \ \forall t \leqslant s$$

成立．因此，将上式在区间 $[t, s]$ 上积分，考虑到 $u(s, s)=1$，则

$$u(t, s)e^{k(t-s)} \leqslant 1$$

由此

$$\|\boldsymbol{\Phi}(t, s)\| \leqslant u(t, s) \leqslant e^{-k(t-s)}, \quad \forall t \leqslant s$$

再加上 $t \geqslant s$ 情况下的结果，则得

$$\|\boldsymbol{\Phi}(t, s)\| \leqslant e^{k|t-s|} \quad \forall t, s \in R.$$

(ii) 设 $\boldsymbol{S}$ 为任意实对称矩阵，α 为正实数，则有下列性质．

(a) $\boldsymbol{S} \geqslant \alpha \boldsymbol{I} \Rightarrow \|\boldsymbol{S}\| \geqslant \alpha$

(b) $\|\boldsymbol{S}\| \leqslant \alpha \Rightarrow \boldsymbol{S} \leqslant \alpha \boldsymbol{I}$.

(证明) 根据 B-I-11 中定理 12，则

$$\lambda_i(\boldsymbol{S}) \leqslant |\lambda_i(\boldsymbol{S})| \leqslant \|\boldsymbol{S}\|$$

成立．于是，若 $\boldsymbol{S} \geqslant \alpha \boldsymbol{I}$，由 B-I-11 中系 19，则

$$\|\boldsymbol{S}\| \geqslant \lambda_i(\boldsymbol{S}) \geqslant \alpha$$

因此，(a) 得到证明．若 $\|\boldsymbol{S}\| \leqslant \alpha$，则

$$\lambda_i(\boldsymbol{S}) \leqslant \|\boldsymbol{S}\| \leqslant \alpha$$

所以，由 B-I-11 中系 19，$(\alpha \boldsymbol{I} - \boldsymbol{S})$ 是准正定的，因此，(b) 得到证明．

定秩系统

在 A-I-8 中，作为线性变系数系统的特殊情况，曾定义过定秩系统．关于可控性及可观测性，定秩系统和常系数系统具有非常相近的性质．例如，和一般变系数系统不同，既使所谓的状态转移矩阵 $\boldsymbol{\Phi}(t, t_0)$ 都不知道，也可以直接由系数矩阵判定其可控性(可观测性)．下面，与以上讲过的关于常系数系统的结果相对比，来讲述一下定秩系统的可控性、可观测性方面的性质．其证明等例如可参考文献[81]．

那么，和以往一样，假定变系数系统

$$\begin{aligned} \dot{\boldsymbol{x}}(t) &= \boldsymbol{A}(t)\boldsymbol{x}(t) + \boldsymbol{B}(t)\boldsymbol{u}(t), \\ & \qquad \boldsymbol{A}(t) \in R^{n\times n}, \ \boldsymbol{B}(t) \in R^{n\times r} \\ \boldsymbol{y}(t) &= \boldsymbol{C}(t)\boldsymbol{x}(t), \qquad \boldsymbol{C}(t) \in R^{m\times n} \end{aligned}$$

表示成 $(\boldsymbol{A}(\cdot), \boldsymbol{B}(\cdot), \boldsymbol{C}(\cdot))$．现在，只要 $\boldsymbol{A}(\cdot)$, $\boldsymbol{B}(\cdot)$, $\boldsymbol{C}(\cdot)$ 连续可微，对于由它们组成的两个矩阵 $\boldsymbol{Q}_k \boldsymbol{R}_k$.

$$\boldsymbol{Q}_k(t) \triangleq [\boldsymbol{P}_0(t), \boldsymbol{P}_1(t), \cdots, \boldsymbol{P}_{k-1}(t)],$$

式中

$$\begin{aligned} \boldsymbol{P}_j(t) &= -\boldsymbol{A}(t)\boldsymbol{P}_{j-1}(t) + \dot{\boldsymbol{P}}_{j-1}(t), \\ \boldsymbol{P}_0 &\triangleq \boldsymbol{B}(t), \\ \boldsymbol{R}_k^T(t) &\triangleq [\boldsymbol{S}_0^T(t), \boldsymbol{S}_1^T(t), \cdots, \boldsymbol{S}_{k-1}^T(t)], \end{aligned}$$

式中

$$\begin{aligned} \boldsymbol{S}_k(t) &= \boldsymbol{S}_{k-1}(t)\boldsymbol{A}(t) + \dot{\boldsymbol{S}}_{k-1}(t), \\ \boldsymbol{S}_0(t) &\triangleq \boldsymbol{C}(t), \end{aligned}$$

则存在某正整数 α, β, 当 $\operatorname{rank} \boldsymbol{Q}_{\alpha}(t)=\operatorname{rank} \boldsymbol{Q}_{\alpha+1}(t)=q_c(\leqslant n)$, $\forall t \in R$, $\operatorname{rank} \boldsymbol{R}_{\beta}(t)=\operatorname{rank} \boldsymbol{R}_{\beta+1}(t)=q_0(\leqslant n)$, $\forall t \in R$ 成立时, $(\boldsymbol{A}(\cdot), \boldsymbol{B}(\cdot), \boldsymbol{C}(\cdot))$称为定秩系统. 一般, 定秩系统的条件是相当严格的, 但是作为其特殊情况, 当然也包括常系数系统在系数矩阵的元素中具有解析函数的系统.

对于定秩系统, 因为在某区间$[t_0, t_1]$上完全可控与在任意区间上完全可控是等价的, 则与常系数情况一样, 谈到可控性我们将只说"是可控的".

可控性, 可观测性的充分必要条件可以分别用 $q_c=n$, $q_0=n$ 给出. 一般, 在定秩系统中, $\operatorname{rank} \boldsymbol{R}_{\beta}(t) \boldsymbol{Q}_{\alpha}(t)$ 为一定值 $q(\leqslant n)$, 而 $q=n$ 是可控且可观测的充分必要条件.

常系数系统中的典范结构定理及由其得到的系 6, 7 等, 在定秩系统中照样成立. 亦即, 当设 $\operatorname{rank} \boldsymbol{R}_{\beta}(t) \boldsymbol{Q}_{\alpha}(t)=q$ 时, 则有与$(\boldsymbol{A}(\cdot), \boldsymbol{B}(\cdot), \boldsymbol{C}(\cdot))$具有相同脉冲响应的 q 维定秩系统$(\boldsymbol{A}_1(\cdot), \overline{\boldsymbol{B}}_1(\cdot), \overline{\boldsymbol{C}}_1(\cdot))$存在. 因此, $(\boldsymbol{A}(\cdot), \boldsymbol{B}(\cdot), \boldsymbol{C}(\cdot))$为其脉冲响应最小实现的充分必要条件是, $(\boldsymbol{A}(\cdot), \boldsymbol{B}(\cdot), \boldsymbol{C}(\cdot))$可控且可观测. 换句话说, $\operatorname{rank} \boldsymbol{R}_{\beta}(t) \boldsymbol{Q}_{\alpha}(t)=n$. 而且, 若两个定秩系统$(\boldsymbol{A}(\cdot), \boldsymbol{B}(\cdot), \boldsymbol{C}(\cdot))$和$(\overline{\boldsymbol{A}}(\cdot), \overline{\boldsymbol{B}}(\cdot), \boldsymbol{C}(\cdot))$为同一脉冲响应的最小实现, 则其相互代数等价(A-I-7)的结果也成立.

A-I-12　可控系统的典范形

考虑常系数系统

$$\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{u},\ \ \boldsymbol{y}=\boldsymbol{C}\boldsymbol{x}+\boldsymbol{D}\boldsymbol{u} \tag{1}$$

$$\boldsymbol{x}(t)\in R^n,\ \ \boldsymbol{y}(t)\in R^m,\ \ \boldsymbol{u}(t)\in R^r$$

由某个 $n\times n$ 正则矩阵 $\boldsymbol{T}$ 进行座标变换

$$\boldsymbol{x}=\boldsymbol{T}\tilde{\boldsymbol{x}}\quad \det \boldsymbol{T}\neq 0 \tag{2}$$

得到和它代数等价的常系数系统

$$\dot{\tilde{\boldsymbol{x}}}=\tilde{\boldsymbol{A}}\tilde{\boldsymbol{x}}+\tilde{\boldsymbol{B}}\boldsymbol{u},\ \ \boldsymbol{y}=\tilde{\boldsymbol{C}}\boldsymbol{x}+\tilde{\boldsymbol{D}}\boldsymbol{u}. \tag{1}$$

两者系数矩阵之间存在下列关系(参看 A-I-10)

$$\tilde{\boldsymbol{A}}=\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T},\ \tilde{\boldsymbol{B}}=\boldsymbol{T}^{-1}\boldsymbol{B},\ \tilde{\boldsymbol{C}}=\boldsymbol{C}\boldsymbol{T},\ \tilde{\boldsymbol{D}}=\boldsymbol{D}. \tag{3}$$

对于这两个代数等价的系统，下列性质成立:

a) 可控性，可观测性保存[1];

b) 特征值保存，即 $\Lambda(\boldsymbol{A})=\Lambda(\tilde{\boldsymbol{A}})$;

c) 输入输出关系(脉冲响应矩阵，传递函数矩阵)相同.

结合着上述性质考虑，选择适当的矩阵 $\boldsymbol{T}$，使方程式成为尽可能简单的形式，这种形式称为典范形(canonical form). 由于不可能同时使 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$ 全部都变成简单形式，故这里只讨论 $\boldsymbol{A}$, $\boldsymbol{B}$ 的典范形，以下只限于可控系统. 此外，在可观测系统中，为了求 $\boldsymbol{C}$, $\boldsymbol{A}$ 的简单形，在以下的叙述中将 $\boldsymbol{A}$ 换成 $\boldsymbol{A}^T$, $\boldsymbol{B}$ 换成 $\boldsymbol{C}^T$ 即可. 根据对偶性，这是显然的.

开始先讨论单输入系统(即 $r=1$ 的情况)，而后再推广到一般的多输入系统($r\geqslant 1$).

单输入系统

现来求可控单输入系统

$$\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}u \tag{4}$$

的几个典范形. 首先，选取

$$\boldsymbol{T}_1\triangleq[\boldsymbol{b},\ \boldsymbol{A}\boldsymbol{b},\ \cdots,\ \boldsymbol{A}^{n-1}\boldsymbol{b}] \tag{5}$$

作为变换矩阵 $\boldsymbol{T}$. 因(4)式可控，则 $\det \boldsymbol{T}_1\neq 0$. 因 $\boldsymbol{b}$ 包含在 $\boldsymbol{T}_1$ 的第 1 列，可见，$\boldsymbol{b}=\boldsymbol{T}_1\boldsymbol{e}_1$，即

$$\boldsymbol{T}_1^{-1}\boldsymbol{b}=\boldsymbol{e}_1. \tag{6}$$

其次，矩阵 $\boldsymbol{A}\boldsymbol{T}_1$ 的第 1 列是 $\boldsymbol{A}\boldsymbol{b}$，它等于 $\boldsymbol{T}_1$ 的第 2 列，即等于 $\boldsymbol{T}_1\boldsymbol{e}_2$. 以下同样，得

$$\boldsymbol{A}\boldsymbol{T}_1=[\boldsymbol{T}_1\boldsymbol{e}_2,\ \boldsymbol{T}_1\boldsymbol{e}_3,\ \cdots,\ \boldsymbol{T}_1\boldsymbol{e}_n,\ \boldsymbol{A}^n\boldsymbol{b}] \tag{7}$$

这里假定 $\boldsymbol{A}$ 的特征多项式，即 $\det(s\boldsymbol{I}-\boldsymbol{A})$ 为

$$\det(s\boldsymbol{I}-\boldsymbol{A})=s^n+a_1s^{n-1}+\cdots+a_{n-1}s+a_n=\psi(s) \tag{8}$$

1) 详细地讲，当且仅当其中一个系统是可控(可观测)时，则另一个系统也可控(可观测).

根据凯莱-哈密顿定理

$$\psi(\boldsymbol{A})=\boldsymbol{A}^{n}+a_{1}\boldsymbol{A}^{n-1}+\cdots+a_{n-1}\boldsymbol{A}+a_{n}\boldsymbol{I}=\boldsymbol{0}\tag{9}$$

成立、因此，得

$$\begin{aligned}\boldsymbol{A}^{n}\boldsymbol{b}&=-a_{1}\boldsymbol{A}^{n-1}\boldsymbol{b}-\cdots-a_{n-1}\boldsymbol{A}\boldsymbol{b}-a_{n}\boldsymbol{b}\\&=-a_{1}\boldsymbol{T}_{1}\boldsymbol{e}_{n}-\cdots-a_{n-1}\boldsymbol{T}_{1}\boldsymbol{e}_{2}-a_{n}\boldsymbol{T}_{1}\boldsymbol{e}_{1}\\&=\boldsymbol{T}_{1}\sum_{j=1}^{n}(-a_{n-j+1})\boldsymbol{e}_{j}\end{aligned}\tag{10}$$

代入(7)式，得

$$\boldsymbol{A}\boldsymbol{T}_{1}=\boldsymbol{T}_{1}\left[\boldsymbol{e}_{2},\ \boldsymbol{e}_{3},\ \cdots,\ \boldsymbol{e}_{n}\sum_{j=1}^{n}(-a_{n-j+1})\boldsymbol{e}_{j}\right]\tag{11}$$

由(6)，(11)式可以得出第1个典范形$(\boldsymbol{A}_{1}^{c},\ \boldsymbol{b}^{c})$

$$(\boldsymbol{A},\ \boldsymbol{b})\xrightarrow{\boldsymbol{T}_{1}}(\boldsymbol{A}_{1}^{c},\ \boldsymbol{b}_{1}^{c})\tag{12}$$

式中

$$\boldsymbol{A}_{1}^{c}=\begin{bmatrix}0 & & & -a_{n}\\ 1 & \ddots & \boldsymbol{0} & -a_{n-1}\\ & \ddots & \ddots & \vdots\\ \boldsymbol{0} & & 0 & -a_{2}\\ & & 1 & -a_{1}\end{bmatrix},\quad \boldsymbol{b}_{1}^{c}=\begin{bmatrix}1\\0\\ \vdots\\ \vdots\\0\end{bmatrix}\tag{13}$$

其次，设 $\boldsymbol{P}$ 为顺列矩阵

$$\boldsymbol{P}\triangleq[\boldsymbol{e}_{n},\ \boldsymbol{e}_{n-1},\ \cdots,\ \boldsymbol{e}_{1}]\tag{14}$$

及

$$\boldsymbol{T}_{2}\triangleq\boldsymbol{T}_{1}\boldsymbol{P}=[\boldsymbol{A}^{n-1}\boldsymbol{b},\ \boldsymbol{A}^{n-2}\boldsymbol{b},\ \cdots,\ \boldsymbol{b}]。\tag{15}$$

用和推导(6)，(11)式完全相同的思路，可以得到第2个典范形

$$(\boldsymbol{A},\ \boldsymbol{b})\xrightarrow{\boldsymbol{T}_{2}}(\boldsymbol{A}_{2}^{c},\ \boldsymbol{b}_{2}^{c})\tag{16}$$

式中

$$\boldsymbol{A}_{2}^{c}=\begin{bmatrix}-a_{1} & 1 & & \\ -a_{2} & 0 & \ddots & \boldsymbol{0}\\ \vdots & & \ddots & \ddots\\ \vdots & \boldsymbol{0} & \ddots & 1\\ -a_{n} & & & 0\end{bmatrix},\quad \boldsymbol{b}_{2}^{c}=\begin{bmatrix}0\\ \vdots\\ \vdots\\0\\1\end{bmatrix}\tag{17}$$

而且，设矩阵 $\boldsymbol{Q}$ 为

$$\boldsymbol{Q}\triangleq\begin{bmatrix}1 & & & \\ a_{1} & \ddots & \boldsymbol{0} & \\ a_{2} & \ddots & \ddots & \\ & \ddots & \ddots & \ddots \\ a_{n-1} & \cdots\ a_{2} & a_{1} & 1\end{bmatrix}\tag{18}$$

定义 $\boldsymbol{T}_{3}$ 为

$$\boldsymbol{T}_3 \triangleq \boldsymbol{T}_2\boldsymbol{Q} = \left[\sum_{j=0}^{n-1} a_j \boldsymbol{A}^{n-j-1}\boldsymbol{b} \;\vdots\; \sum_{j=0}^{n-2} a_j \boldsymbol{A}^{n-j-2}\boldsymbol{b} \;\vdots\; \cdots \;\vdots\; a_0\boldsymbol{A}\boldsymbol{b} + a_1\boldsymbol{b} \;\vdots\; a_0\boldsymbol{b}\right] \tag{19}$$

其中假定 $a_0 \triangleq 1$. 因此, $\boldsymbol{b} = a_0\boldsymbol{b} = \boldsymbol{T}_3\boldsymbol{e}_n$. 由(19)式, $\boldsymbol{A}\boldsymbol{T}_3$ 的第 1 列是 $\sum_{j=0}^{n-1} a_j \boldsymbol{A}^{n-j}\boldsymbol{b}$, 利用(9)式，它等于 $-a_n\boldsymbol{b}$. 因 $\boldsymbol{T}_3$ 的第 n 列是 $a_0\boldsymbol{b} = \boldsymbol{b}$, 则 $-a_n\boldsymbol{b} = -a_n\boldsymbol{T}_3\boldsymbol{e}_n$. 其次，对于 $\boldsymbol{A}\boldsymbol{T}_3$ 的第 2 列，由(19)式，

$$\begin{aligned} \boldsymbol{A}\boldsymbol{T}_3 \text{ 的第 2 列} &= \sum_{j=0}^{n-2} a_j \boldsymbol{A}^{n-j+1}\boldsymbol{b} = \sum_{j=0}^{n-1} a_j \boldsymbol{A}^{n-j-1}\boldsymbol{b} - a_{n-1}\boldsymbol{b} \\ &= \boldsymbol{T}_3\boldsymbol{e}_1 - a_{n-1}\boldsymbol{T}_3\boldsymbol{e}_n = \boldsymbol{T}_3(\boldsymbol{e}_1 - a_{n-1}\boldsymbol{e}_n) \end{aligned} \tag{20}$$

成立. 以下同样，得

$$\boldsymbol{A}\boldsymbol{T}_3 = \boldsymbol{T}_3[-a_n\boldsymbol{e}_n,\ \boldsymbol{e}_1 - a_{n-1}\boldsymbol{e}_n,\ \cdots,\ \boldsymbol{e}_{n-2} - a_2\boldsymbol{e}_n,\ \boldsymbol{e}_{n-1} - a_1\boldsymbol{e}_n] \tag{21}$$

于是，便得到第 3 典范形

$$(\boldsymbol{A},\ \boldsymbol{b}) \xrightarrow{\boldsymbol{T}_3} (\boldsymbol{A}_3^c,\ \boldsymbol{b}_3^c)$$

式中

$$\boldsymbol{A}_3^c = \begin{bmatrix} 0 & 1 & & & \\ & \ddots & \ddots & \boldsymbol{0} & \\ & \boldsymbol{0} & \ddots & \ddots & \\ & & & 0 & 1 \\ -a_n & -a_{n-1} & \cdots & & -a_1 \end{bmatrix},\quad \boldsymbol{b}_3^c = \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{22}$$

它和在 A-I-11 中讲过的可控典范形一致.

该第 3 典范形也可按如下推导.

设 $\boldsymbol{T}_2^{-1}$ 的第 1 行为 $\boldsymbol{h}$, 即

$$\boldsymbol{h} \triangleq \boldsymbol{e}_1^T \boldsymbol{T}_2^{-1} \tag{23}$$

则由定义得

$$\begin{cases} \boldsymbol{h}\boldsymbol{A}^k\boldsymbol{b} = 0 \\ \boldsymbol{h}\boldsymbol{A}^{n-1}\boldsymbol{b} = 1, \end{cases} \qquad 0 \leqslant k < n-1 \tag{24}$$

设矩阵 $\boldsymbol{S}$ 为

$$\boldsymbol{S} \triangleq \begin{bmatrix} \boldsymbol{h} \\ \boldsymbol{h}\boldsymbol{A} \\ \vdots \\ \boldsymbol{h}\boldsymbol{A}^{n-1} \end{bmatrix} \tag{25}$$

利用(24)式可以证明, $\boldsymbol{S}$ 的行线性独立，即 $\boldsymbol{S}$ 是正则的. 对 $\boldsymbol{S}\boldsymbol{A}$ 的各行进行和由(5)式得出(11)式时同样的论证，得

$$\boldsymbol{S}\boldsymbol{A} = \begin{bmatrix} \boldsymbol{e}_2^T \\ \boldsymbol{e}_3^T \\ \vdots \\ \boldsymbol{e}_n^T \\ \sum_{j=1}^{n} -a_{n-j+1}\boldsymbol{e}_j^T \end{bmatrix} \boldsymbol{S} \tag{26}$$

因此

$$\boldsymbol{S}\boldsymbol{A}\boldsymbol{S}^{-1}=\boldsymbol{A}_3^c \tag{27}$$

再利用(24)式，显然

$$\boldsymbol{S}\boldsymbol{b}=\boldsymbol{e}_n=\boldsymbol{b}_3^c \tag{28}$$

因此

$$(\boldsymbol{A},\ \boldsymbol{b})\xrightarrow{\boldsymbol{S}^{-1}}(\boldsymbol{A}_3^c,\ \boldsymbol{b}_3^c) \tag{29}$$

而由(23)和(25)式，$\boldsymbol{S}$ 也可以表示如下：

$$\boldsymbol{S}=\begin{bmatrix} \boldsymbol{e}_1^T\boldsymbol{T}_2^{-1} \\ \boldsymbol{e}_1^T\boldsymbol{T}_2^{-1}\boldsymbol{A} \\ \vdots \\ \boldsymbol{e}_1^T\boldsymbol{T}_2^{-1}\boldsymbol{A}^{n-1} \end{bmatrix}=\begin{bmatrix} \boldsymbol{e}_1^T \\ \boldsymbol{e}_1^T\boldsymbol{A}_2^c \\ \vdots \\ \boldsymbol{e}_1^T(\boldsymbol{A}_2^c)^{n-1} \end{bmatrix}\boldsymbol{T}_2^{-1} \tag{30}$$

这里若设

$$\boldsymbol{R}\triangleq\begin{bmatrix} \boldsymbol{e}_1^T \\ \boldsymbol{e}_1^T\boldsymbol{A}_2^c \\ \vdots \\ \boldsymbol{e}_1^T(\boldsymbol{A}_2^c)^{n-1} \end{bmatrix} \tag{31}$$

试乘一下就很容易证明，$\boldsymbol{R}=\boldsymbol{Q}^{-1}$.

在许多文献中都有上述典范形的推导[84-87].

在这里假定，$\boldsymbol{A}$ 具有 n 个相异的特征值 $\lambda_1,\ \cdots,\ \lambda_n$，则 $\boldsymbol{A}$ 具有 n 个线性独立的特征向量 $\boldsymbol{x}_1,\ \cdots,\ \boldsymbol{x}_n$，$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$ 可以由 $\boldsymbol{T}\triangleq[\boldsymbol{x}_1,\ \cdots,\ \boldsymbol{x}_n]$ 变换成以 $\boldsymbol{A}$ 的特征值为主对角线上元素的对角矩阵 $\boldsymbol{\Lambda}$（参看 B-I-11）. 现在我们研究该对角形与上述第 3 典范形的关系. 因为假定 $\boldsymbol{A}$ 具有 n 个相异的特征值，则范得蒙(van der Monde)矩阵

$$\boldsymbol{V}\triangleq\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_n^2 \\ \vdots & \vdots & & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{bmatrix} \tag{32}$$

是正则的（参看 P-I-1 中第 26 页）. 因此，设

$$\boldsymbol{T}_4\triangleq\boldsymbol{T}_3\boldsymbol{V}. \tag{33}$$

由(21)，(32)，(33)式，$\boldsymbol{A}\boldsymbol{T}_4$ 的第 j 列为

$$\begin{aligned}\boldsymbol{A}\boldsymbol{T}_4\text{ 的第 } j \text{ 列}=\boldsymbol{T}_3\{&-a_n\boldsymbol{e}_n\times1+(\boldsymbol{e}_1-a_{n-1}\boldsymbol{e}_n)\times\lambda_j \\ &+(\boldsymbol{e}_2-a_{n-2}\boldsymbol{e}_n)\times\lambda_j^2+\cdots+(\boldsymbol{e}_{n-2}-a_2\boldsymbol{e}_n)\times\lambda_j^{n-2} \\ &+(\boldsymbol{e}_{n-1}-a_1\boldsymbol{e}_n)\times\lambda_j^{n-1}\}\end{aligned} \tag{34}$$

因 λ_j 是 $\boldsymbol{A}$ 的特征值，则 $\psi(\lambda_j)=0$. 因此

$$-a_n\boldsymbol{e}_n\times1-a_{n-1}\boldsymbol{e}_n\lambda_j-a_{n-2}\boldsymbol{e}_n\lambda_j^2-\cdots-a_2\boldsymbol{e}_n\lambda_j^{n-2}-a_1\boldsymbol{e}_n\lambda_j^{n-1}=\boldsymbol{e}_n\lambda_j^n \tag{35}$$

将上式代入(34)式，得

$$\boldsymbol{A}\boldsymbol{T}_4\text{ 的第 } j \text{ 列}=\boldsymbol{T}_3\{\boldsymbol{e}_1\lambda_j+\boldsymbol{e}_2\lambda_j^2+\cdots+\boldsymbol{e}_{n-2}\lambda_j^{n-2}+\boldsymbol{e}_{n-1}\lambda_j^{n-1}+\boldsymbol{e}_n\lambda_j^n\} \tag{36}$$

右边｛　｝中等于 $\lambda_j\times$（$\boldsymbol{V}$ 的第 j 列），则

$$\boldsymbol{A}\boldsymbol{T}_4\text{ 的第 } j \text{ 列}=\boldsymbol{T}_3(\lambda_j\boldsymbol{V}\boldsymbol{e}_j)=\boldsymbol{T}_3\boldsymbol{V}(\lambda_j\boldsymbol{e}_j)=\boldsymbol{T}_4(\lambda_j\boldsymbol{e}_j) \tag{37}$$

因而

$$AT_4 = T_4[\lambda_1 e_1,\ \lambda_2 e_2,\ \cdots,\ \lambda_n e_n] \tag{38}$$

于是可以看出，由变换 T_4 可以得到对角形 Λ

$$(A,\ b) \xrightarrow{T_4} (A_4^c,\ b_4^c) \tag{39}$$

式中

$$A_4^c = \Lambda \tag{40}$$

仍由(37)式

$$AT_4 \text{ 的第 } j \text{ 列} = T_3 V \lambda_j e_j$$

将上式两边左乘以 T_3^{-1}，利用(33)式，得

$$(T_3^{-1}AT)(V \text{ 的第 } j \text{ 列}) = \lambda_j (V \text{ 的第 } j \text{ 列})$$

由此可见，V 的第 j 列，即$[1,\ \lambda_j,\ \lambda_j^2,\ \cdots,\ \lambda_j^{n-1}]^T$，是与 $A_3^c = T_3^{-1}AT_3$ 的 λ_j 对应的特征向量.

$$b_4^c = V^{-1}e_n \tag{41}$$

以上单输入情况下的标准形，除了 A_4^c 中特征值的序号以外，对于所有给定的 $A,\ b$ 都唯一确定.

多输入系统

考虑多输入系统(1)，且假定 B 具有最大秩 r. 将 B 的第 j 列写成 b_j.

设 $\sigma_j(j=1,\ \cdots,\ r)$ 为满足下列三个条件

(i)
$$\sigma_j \geqslant 1 \quad \forall j \tag{42}$$

(ii)
$$\sum_{j=1}^{r} \sigma_j = n \tag{43}$$

(iii)
$$\begin{aligned} T_5 \triangleq [&b_1,\ Ab_1,\ \cdots,\ A^{\sigma_1-1}b_1 \mid b_2,\ Ab_2,\ \cdots,\ A^{\sigma_2-1}b_2 \mid \cdots \\ &\mid b_r,\ Ab_r,\ \cdots,\ A^{\sigma_r-1}b_r],\ \det T_5 \neq 0 \end{aligned} \tag{44}$$

的整数组. 显然，若可控且 $\operatorname{rank} B = r$，则至少有一组满足(i)，(ii)，(iii)的 σ_j 存在.

和(6)式一样，可得到下式

$$T_5^{-1}B = [e_{d_1+1},\ e_{d_2+1},\ \cdots,\ e_{d_r+1}] \tag{45}$$

其中设

$$d_1 \triangleq 0,\ d_i \triangleq \sum_{j=1}^{i-1} \sigma_j\ (i \geqslant 2),\ d_{r+1} \triangleq n \tag{46}$$

而且，和(7)式一样

$$\begin{aligned} AT_5 = [&T_5 e_{d_1+2},\ T_5 e_{d_1+3},\ \cdots,\ T_5 e_{d_2} A^{\sigma_1} b_1 \mid T_5 e_{d_2+2},\ T_5 e_{d_2+3},\ \cdots, \\ &T_5 e_{d_3} A^{\sigma_2} b_2 \mid \cdots \mid T_5 e_{d_r+2},\ T_5 e_{d_r+3},\ \cdots,\ T_5 e_n A^{\sigma_r} b_r] \end{aligned} \tag{47}$$

在这里，虽然不知道 $A^{\sigma_i}b_i$ 是什么样的向量，但因 T_5 是正则的，则确实可用 T_5 列的线性组合表示. 即有满足

$$A^{\sigma_i}b_i = -\sum_{j=1}^{r}\sum_{k=1}^{\sigma_j} \alpha_k^{ji} A^{k-1}b_j = -T_5 \sum_{j=1}^{r}\sum_{k=1}^{\sigma_j} \alpha_k^{ji} e_{d_j+k} \tag{48}$$

的 $\alpha_k^{ji}(i,\ j=1,\ \cdots,\ r;\ k=1,\ \cdots,\ \sigma_j)$ 存在.

由(45)，(47)，(48)式可以得到相当于(12)，(13)式的第 1 典范形

$$(\boldsymbol{A},\ \boldsymbol{B}) \xrightarrow{\boldsymbol{T}_5} (\boldsymbol{A}_5^c,\ \boldsymbol{B}_5^c) \tag{49}$$

式中

$$\boldsymbol{A}_5^c = \left[\begin{array}{c|c|c} \boldsymbol{A}_{11} & \cdots\cdots & \boldsymbol{A}_{1r} \\ \hline \vdots & & \vdots \\ \hline \boldsymbol{A}_{r1} & \cdots\cdots & \boldsymbol{A}_{rr} \end{array}\right],$$

$$\boldsymbol{A}_{ii} = \left[\begin{array}{cccc:c} 0 & & & \boldsymbol{0} & -\alpha_1^{ii} \\ 1 & \ddots & & & \vdots \\ & \ddots & \ddots & 0 & \\ \boldsymbol{0} & & \ddots & 1 & -\alpha_{\sigma_i}^{ii} \end{array}\right] = \sigma_i \times \sigma_i \quad i=1,\ \cdots,\ r \tag{50}$$

$$\boldsymbol{A}_{ji} = \left[\begin{array}{c:c} & -\alpha_1^{ji} \\ \boldsymbol{0} & \vdots \\ & -\alpha_{\sigma_j}^{ji} \end{array}\right] = \sigma_j \times \sigma_i \quad i,\ j=1,\ \cdots,\ r;\ i \neq j$$

$$\boldsymbol{B}_5^c = [\boldsymbol{e}_{d_1+1},\ \boldsymbol{e}_{d_2+1},\ \cdots,\ \boldsymbol{e}_{d_r+1}]$$

因为σ_j在满足条件(i)—(iii)之下是任意的，则可能有各种组合，所以α_k^{ji}值也是变的. 因此，和单输入的情况不同，即使给出$\boldsymbol{A}$, $\boldsymbol{B}$, 典范形$\boldsymbol{A}_5^c$也不是唯一确定的. 亦即，虽然具有(50)式的结构，但是子矩阵的阶数与非零元素α_k^{ij}的值不同，可能存在两个以上的这种典范形.

其次，设顺列矩阵$\boldsymbol{P}$为

$$\boldsymbol{P} \triangleq \boldsymbol{P}_1 \dotplus \cdots \dotplus \boldsymbol{P}_r \tag{51}$$

$$\boldsymbol{P}_i = \begin{bmatrix} \boldsymbol{0} & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & \boldsymbol{0} \end{bmatrix} = \sigma_i \times \sigma_i \quad i=1,\ \cdots,\ r \tag{52}$$

若将$\boldsymbol{T}_6$定义成

$$\boldsymbol{T}_6 \triangleq \boldsymbol{T}_5\boldsymbol{P} = [\boldsymbol{A}^{\sigma_1-1}\boldsymbol{b}_1,\ \cdots,\ \boldsymbol{b}_1 \mid \boldsymbol{A}^{\sigma_2-1}\boldsymbol{b}_2,\ \cdots,\ \boldsymbol{b}_2 \mid \cdots \mid \boldsymbol{A}^{\sigma_r-1}\boldsymbol{b}_r,\ \cdots,\ \boldsymbol{b}_r] \tag{53}$$

则可得第2典范形

$$(\boldsymbol{A},\ \boldsymbol{B}) \xrightarrow{\boldsymbol{T}_6} (\boldsymbol{A}_6^c,\ \boldsymbol{B}_6^c) \tag{54}$$

式中

$$\boldsymbol{A}_6^c = \left[\begin{array}{c|c|c} \boldsymbol{A}'_{11} & \cdots\cdots & \boldsymbol{A}'_{1r} \\ \hline \vdots & & \vdots \\ \hline \boldsymbol{A}'_{r1} & \cdots\cdots & \boldsymbol{A}'_{rr} \end{array}\right],$$

$$\boldsymbol{A}'_{ii} = \left[\begin{array}{c:cccc} -\alpha_{\sigma_i}^{ii} & 1 & & & \boldsymbol{0} \\ & 0 & \ddots & & \\ \vdots & & \ddots & \ddots & 1 \\ -\alpha_1^{ii} & & \boldsymbol{0} & \ddots & 0 \end{array}\right], \tag{55}$$

$$
A'_{j_i}=\left[\begin{array}{c:c} -\alpha_{\sigma_j}^{j_i} & \\ \vdots & \mathbf{0} \\ -\alpha_1^{j_i} & \end{array}\right]
$$

$$
B_6^c=[e_{d_1},\ e_{d_2},\ \cdots,\ e_{d_r},\ e_n]
$$

前面讲过两种求单输入第 3 典范形的方法，将其中任何一种方法加以推广，即可求出在多输入情况下的第 3 典范形．开始我们先将使用 S 矩阵的方法予以推广．作 T_6^{-1}，定义行向量 h_i 为

$$
h_i \triangleq e_{d_i+1}^T T_6^{-1} \quad i=1,\ \cdots,\ r \tag{56}
$$

和单输入的情况一样，立即可以看到

$$
h_i A^k b_j=\begin{cases} 0 & 0\leqslant k\leqslant \sigma_j-1 \\ \delta_{ij} & k=\sigma_j-1 \\ * & \sigma_j\leqslant k \end{cases} \qquad * \text{ 表示不一定为 0 的数} \tag{57}
$$

其次，设矩阵 S 为

$$
S \triangleq \left[\begin{array}{c} h_1 \\ h_1 A \\ \vdots \\ h_1 A^{\sigma_1-1} \\ \hline h_2 \\ h_2 A \\ \vdots \\ h_2 A^{\sigma_2-1} \\ \hdashline \vdots \\ \hdashline h_r \\ h_r A \\ \vdots \\ h_r A^{\sigma_r-1} \end{array}\right]. \tag{58}
$$

和前面一样，容易证明 S 是正则的．因此，考虑变换

$$
(A,\ B)\xrightarrow{S^{-1}}(SAS^{-1},\ SB). \tag{59}
$$

首先，SA 的第 1 行是 h_1A，它等于 S 的第 2 行，即等于 e_2^TS．第 2 行是 $h_1A^2=e_3^TS$ 等等．于是可以得到相当于(47)式的关系式

$$
SA=\left[\begin{array}{c} e_{d_1+2}^T S \\ e_{d_1+3}^T S \\ \vdots \\ e_{d_2}^T S \\ h_1 A^{\sigma_1} \\ \hdashline \vdots \\ \hdashline e_{d_r+2}^T S \\ \vdots \\ e_n^T S \\ h_r A^{\sigma_r} \end{array}\right] \tag{60}
$$

第 d_i 行是 $\boldsymbol{h}_i\boldsymbol{A}^{\sigma_i}$, 虽然不知道它是什么样的向量,但是因 $\boldsymbol{S}$ 是正则的, 所以可以用和推导(48)式相同的方法证明,有满足

$$\boldsymbol{h}_i\boldsymbol{A}^{\sigma_i}=-\sum_{j=1}^{r}\sum_{k=1}^{\sigma_j}\beta_k^{ji}\boldsymbol{h}_j\boldsymbol{A}^{k-1}=-\sum_{j=1}^{r}\sum_{k=1}^{\sigma_j}\beta_k^{ji}\boldsymbol{e}_{d_j+k}^T\boldsymbol{S} \tag{61}$$

的 $\beta_k^{ji}(i,\ j=1,\ \cdots,\ r;\ k=1,\ \cdots,\ \sigma_j)$ 存在. 由这些式子可以得到对于 $\boldsymbol{T}_7\triangleq\boldsymbol{S}^{-1}$ 的第 3 典范形

$$(\boldsymbol{A},\ \boldsymbol{B})\xrightarrow{\boldsymbol{T}_7}(\boldsymbol{A}_7^c,\ \boldsymbol{B}_7^c) \tag{62}$$

$$\boldsymbol{A}_7^c=\begin{bmatrix}\boldsymbol{A}_{11}'' & \cdots & \boldsymbol{A}_{1r}''\\ \vdots & & \vdots\\ \boldsymbol{A}_{r1}'' & \cdots & \boldsymbol{A}_{rr}''\end{bmatrix},$$

$$\boldsymbol{A}_{ii}''=\begin{bmatrix}0 & 1 & & \boldsymbol{0}\\ & \ddots & \ddots & \\ & \boldsymbol{0} & 0 & 1\\ -\beta_1^{ii} & \cdots & & -\beta_{\sigma_i}^{ii}\end{bmatrix}$$

$$i=1,\ \cdots,\ r$$

$$\boldsymbol{A}_{ij}''=\begin{bmatrix} & \boldsymbol{0} & \\ \cdots & \cdots & \cdots\\ -\beta_1^{ij} & \cdots\cdots & -\beta_{\sigma_j}^{ij}\end{bmatrix}$$

$$i,\ j=1,\ \cdots,\ r;\ i\neq j \tag{63}$$

要注意,这里和单输入的情况不同

(a) α_k^{ij}, β_k^{ij} 不一定与特征多项式的系数一致;

(b) $\boldsymbol{B}_7^c$ 不一定象(22)式 $\boldsymbol{b}_3^c$ 那样具有只以 0 和 1 为元素的简单形式.

和单输入的情况一样, (58)式的矩阵 $\boldsymbol{S}$ 也可以表示如下:

$$\boldsymbol{S}=\boldsymbol{R}\boldsymbol{T}_6^{-1}$$

$$\boldsymbol{R}\triangleq\begin{bmatrix}\boldsymbol{e}_{d_1+1}^T\\ \boldsymbol{e}_{d_1+1}^T\boldsymbol{A}_6^c\\ \vdots\\ \boldsymbol{e}_{d_1+1}^T(\boldsymbol{A}_6^c)^{\sigma_1-1}\\ \cdots\cdots\cdots\\ \boldsymbol{e}_{d_2+1}^T\\ \boldsymbol{e}_{d_2+1}^T\boldsymbol{A}_6^c\\ \vdots\\ \boldsymbol{e}_{d_2+1}^T(\boldsymbol{A}_6^c)^{\sigma_2-1}\\ \cdots\cdots\cdots\\ \vdots\\ \cdots\cdots\cdots\\ \boldsymbol{e}_n^T\\ \boldsymbol{e}_n^T\boldsymbol{A}_6^c\\ \vdots\\ \boldsymbol{e}_n^T(\boldsymbol{A}_6^c)^{\sigma_r-1}\end{bmatrix}. \tag{64}$$

该 $\boldsymbol{R}$, 因而其逆矩阵 $\boldsymbol{Q}$, 一般也是很复杂的. 但是, 取出 $\boldsymbol{A}_6^c$ 的对角子块 $\boldsymbol{A}'_{ii}$, 作

$$\hat{\boldsymbol{R}}_{ii} \triangleq \begin{bmatrix} \boldsymbol{e}_1^T \\ \boldsymbol{e}_1^T \boldsymbol{A}'_{ii} \\ \vdots \\ \boldsymbol{e}_1^T (\boldsymbol{A}'_{ii})^{\sigma_i - 1} \end{bmatrix} \tag{65}$$

试将 $\hat{\boldsymbol{R}}$ 定义成

$$\hat{\boldsymbol{R}} \triangleq \hat{\boldsymbol{R}}_{11} \dotplus \cdots \dotplus \hat{\boldsymbol{R}}_{rr} \tag{66}$$

这相当于暂且把 $\boldsymbol{A}'_{ii}$ 看成与其它不相干涉, 对它构成和单输入情况下一样的 $\hat{\boldsymbol{R}}_{ii}$, 然后将全体用其直和表示. $\boldsymbol{R}$ 确实是由 $\boldsymbol{A}_6^c$ 得到的, 与此相应, 可以说 $\hat{\boldsymbol{R}}$ 是"强制"地忽略 $\boldsymbol{A}_6^c$ 非主对角线上的元素而构成的. $(\boldsymbol{A}_6^c, \boldsymbol{B}_6^c)$ 由上述 $\hat{\boldsymbol{R}}$ 会变换成什么呢? 我们利用 $\hat{\boldsymbol{R}}$ 的逆矩阵 $\hat{\boldsymbol{Q}} = \hat{\boldsymbol{R}}^{-1}$ 来研究一下. 当然

$$\hat{\boldsymbol{Q}} = \hat{\boldsymbol{Q}}_{11} \dotplus \cdots \dotplus \hat{\boldsymbol{Q}}_{rr} \tag{67}$$

$$\hat{\boldsymbol{Q}}_{ii} = \hat{\boldsymbol{R}}_{ii}^{-1} = \begin{bmatrix} 1 & & & & \\ \alpha_{\sigma_i}^{ii} & \ddots & & \boldsymbol{0} & \\ \alpha_{\sigma_i-1}^{ii} & \ddots & \ddots & & \\ \vdots & & \ddots & \ddots & \\ \alpha_2^{ii} & \cdots & \alpha_{\sigma_i-1}^{ii} & \alpha_{\sigma_i}^{ii} & 1 \end{bmatrix} \tag{68}$$

若令 $\boldsymbol{T}_8 \triangleq \boldsymbol{T}_6 \hat{\boldsymbol{Q}}$, 则显然 $\boldsymbol{T}_8$ 的第 σ_1 列为 $\boldsymbol{b}_1$, 第 $\sigma_1 + \sigma_2 (= d_3)$ 列为 $\boldsymbol{b}_2$, 等等, 由此得

$$\boldsymbol{B} = \boldsymbol{T}_8 [\boldsymbol{e}_{d_2}, \ \boldsymbol{e}_{d_3}, \ \cdots, \ \boldsymbol{e}_{d_r}, \ \boldsymbol{e}_n] \tag{69}$$

则

$$\boldsymbol{B}_8^c \triangleq \boldsymbol{T}_8^{-1} \boldsymbol{B} = [\boldsymbol{e}_{d_2}, \ \boldsymbol{e}_{d_3}, \ \cdots, \ \boldsymbol{e}_{d_r}, \ \boldsymbol{e}_n]. \tag{70}$$

作 $\boldsymbol{A}\boldsymbol{T}_8 = \boldsymbol{A}\boldsymbol{T}_6 \hat{\boldsymbol{Q}}$ 的第 1 列, 得

$$\boldsymbol{A}\boldsymbol{T}_8 \boldsymbol{e}_1 = \boldsymbol{A}^{\sigma_1} \boldsymbol{b}_1 + \alpha_{\sigma_1}^{11} \boldsymbol{A}^{\sigma_1 - 1} \boldsymbol{b}_1 + \cdots + \alpha_2^{11} \boldsymbol{A} \boldsymbol{b}_1 \tag{71}$$

由(48)式得

$$\boldsymbol{A}\boldsymbol{T}_8 \boldsymbol{e}_1 = -\alpha^{11} \boldsymbol{b}_1 - \sum_{j=2}^{r} \sum_{k=1}^{\sigma_j} \alpha_k^{j1} \boldsymbol{A}^{k-1} \boldsymbol{b}_j \tag{72}$$

如已讲过, $\boldsymbol{b}_1$ 是 $\boldsymbol{T}_8$ 的第 σ_1 列, 即 $\boldsymbol{T}_8 \boldsymbol{e}_{\sigma_1}$. $\boldsymbol{A}^{k-1} \boldsymbol{b}_j$ 是 $\boldsymbol{T}_6$(即 $\boldsymbol{T}_8 \hat{\boldsymbol{Q}}^{-1}$)的第 $d_j + k$ 列

$$\boldsymbol{A}^{k-1} \boldsymbol{b}_j = \boldsymbol{T}_8 \hat{\boldsymbol{Q}}^{-1} \boldsymbol{e}_{d_j+k} = \boldsymbol{T}_8 \hat{\boldsymbol{R}} \boldsymbol{e}_{d_j+k} \tag{73}$$

如(66)式, 因 $\hat{\boldsymbol{R}}$ 是子矩阵的直和, 则 $\hat{\boldsymbol{R}}_j \boldsymbol{e}_{d_j+k}$ 是

$$\hat{\boldsymbol{R}} \boldsymbol{e}_{d_j+k} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \hdashline r_1^{jk} \\ \vdots \\ r_{\sigma_j}^{jk} \\ \hdashline 0 \\ \vdots \\ 0 \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}0\\ \vdots\\ 0\end{matrix}}\right\} d_j \text{ 个} \\ \left.\vphantom{\begin{matrix}0\\ \vdots\\ 0\end{matrix}}\right\} \sigma_j \text{ 个} \\ \left.\vphantom{\begin{matrix}0\\ \vdots\\ 0\end{matrix}}\right\} (n - d_{j+1}) \text{ 个} \end{matrix} \tag{74}$$

形的向量．将其代入，得

$$\boldsymbol{A}\boldsymbol{T}_8\boldsymbol{e}_1=\boldsymbol{T}_8\begin{bmatrix}\boldsymbol{0}\\ \vdots\\ 0\\ -\alpha_1^{11}\\ \cdots\cdots\\ \delta_1^{21}\\ \vdots\\ \delta_{\sigma_2}^{21}\\ \cdots\cdots\\ \delta_1^{31}\\ \vdots\\ \delta_{\sigma_3}^{31}\\ \cdots\cdots\\ \vdots\\ \cdots\cdots\\ \delta_1^{r1}\\ \vdots\\ \delta_{\sigma_r}^{r1}\end{bmatrix}\qquad \begin{array}{l}\delta_l^{j1}=-\sum\limits_{k=1}^{\sigma_j}\alpha_k^{11}\gamma_l^{jk}\\ \quad (j=2,\ \cdots,\ r;\ l=1,\ \cdots,\ \sigma_j)\end{array}\tag{75}$$

其次，$\boldsymbol{A}\boldsymbol{T}_8$ 的第 2 列为

$$\begin{aligned}\boldsymbol{A}\boldsymbol{T}_8\boldsymbol{e}_2&=\boldsymbol{A}^{\sigma_1-1}\boldsymbol{b}_1+\alpha_{\sigma_1}^{11}\boldsymbol{A}^{\sigma_1-2}\boldsymbol{b}_1+\cdots+\alpha_3^{11}\boldsymbol{A}\boldsymbol{b}_1\\ &=\boldsymbol{T}_8\boldsymbol{e}_1-\alpha_2^{11}\boldsymbol{b}_1=\boldsymbol{T}_8\boldsymbol{e}_2-\alpha_2^{11}\boldsymbol{T}_8\boldsymbol{e}_{d_1}\end{aligned}\tag{76}$$

反复进行上述计算，最后可以得出

$$(\boldsymbol{A},\ \boldsymbol{B})\xrightarrow{\boldsymbol{T}_8}(\boldsymbol{A}_8^c,\ \boldsymbol{B}_8^c)\tag{77}$$

式中

$$\boldsymbol{A}_8^c=\begin{bmatrix}\widetilde{\boldsymbol{A}}_{11} & \cdots\cdots & \widetilde{\boldsymbol{A}}_{1r}\\ \vdots & & \vdots\\ \widetilde{\boldsymbol{A}}_{r1} & \cdots\cdots & \widetilde{\boldsymbol{A}}_{rr}\end{bmatrix},$$

$$\widetilde{\boldsymbol{A}}_{ii}=\begin{bmatrix}0 & 1 & & \boldsymbol{0}\\ & \ddots & \ddots & \\ \boldsymbol{0} & & 0 & 1\\ -\alpha_1^{ii} & -\alpha_2^{ii} & \cdots & -\alpha_{\sigma_i}^{ii}\end{bmatrix}$$

$$\widetilde{\boldsymbol{A}}_{ji}=\begin{bmatrix}\delta_1^{ji} & \\ \vdots & \boldsymbol{0}\\ \delta_{\sigma_j}^{ji} & \end{bmatrix}\quad (i\neq j)$$

$$\boldsymbol{B}_8^c=\begin{array}{c}\left[\begin{array}{c}\boldsymbol{0}\\ 1\ 0\ 0\cdots 0\\ \hdashline \boldsymbol{0}\\ 0\ 1\ 0\cdots 0\\ \hdashline \vdots\\ \hdashline \boldsymbol{0}\\ 0\quad\cdots 0\ 1\end{array}\right]\end{array}\begin{array}{l}\left.\begin{array}{c}\\ \\ \end{array}\right\}\sigma_1\\ \left.\begin{array}{c}\\ \\ \end{array}\right\}\sigma_2\\ \\ \left.\begin{array}{c}\\ \\ \end{array}\right\}\sigma_r\end{array}\tag{78}$$

这里要注意，它和$(\boldsymbol{A}_7^c, \boldsymbol{B}_7^c)$不同，$\boldsymbol{B}$ 变成很理想的形式，但 $\boldsymbol{A}$ 的非对角子块的形状变了．在文献[87]中使用$(\boldsymbol{A}_8^c, \boldsymbol{B}_8^c)$．

由 $\boldsymbol{T}_5$, $\boldsymbol{T}_6$, $\boldsymbol{T}_7=\boldsymbol{S}^{-1}$, $\boldsymbol{T}_8$ 的作法可以看到，与 $\boldsymbol{b}_1$ 有关的列或行有 σ_1 个，与 $\boldsymbol{b}_2$ 有关的有 σ_2 个…，与 $\boldsymbol{b}_r$ 有关的有 σ_r 个．因此，称 σ_j 为第 j 个系列的长度．而且定义

$$\mu_k \triangleq \text{长度在 } k \text{ 以上的系列数} \tag{79}$$

根据条件(i)，所有系列的长度都在1以上．因此，$\mu_1=r$ 成立．而且显然

$$\mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_\nu > \mu_{\nu+1}=0 \quad \nu \triangleq \max_j \sigma_j \tag{80}$$

变换 $\boldsymbol{T}_5$ 列的顺序，按如下排列作 $\widetilde{\boldsymbol{T}}_5$．

(a) 整个矩阵由 ν 个子块组成，第 ν 个子块包含有各系列中 $\boldsymbol{A}$ 的指数最大的（即 $\boldsymbol{A}^{\sigma_j-1}\boldsymbol{b}_j$, $j=1, \cdots, r$），第$(\nu-1)$个子块包含有指数比前者仅小1的……．当然，第 i 个子块中包含的列数等于 $\mu_{\nu-i+1}$．

(b) 在各个子块中，从左到右，按照由 $\boldsymbol{A}$ 的指数小的到大的排列．对于指数相同的，从左到右，按照由 $\boldsymbol{b}_j$ 下标小的到大的次序排列．例如，在 $n=6$, $r=3$, 设

$$\boldsymbol{T}_5=[\boldsymbol{b}_1 \mid \boldsymbol{b}_2 \ \ \boldsymbol{A}\boldsymbol{b}_2 \ \ \boldsymbol{A}^2\boldsymbol{b}_2 \mid \boldsymbol{b}_3 \ \ \boldsymbol{A}\boldsymbol{b}_3] \tag{81}$$

时，按照(a)，(b)排列，得

$$\widetilde{\boldsymbol{T}}_5=[\boldsymbol{b}_2 \mid \boldsymbol{b}_3 \ \ \boldsymbol{A}\boldsymbol{b}_2 \quad \boldsymbol{b}_1 \ \ \boldsymbol{A}\boldsymbol{b}_3 \ \ \boldsymbol{A}^2\boldsymbol{b}_2]. \tag{82}$$

因 $\widetilde{\boldsymbol{T}}_5$ 是将 $\boldsymbol{T}_5$ 列的顺序交换后得到的，则有某个顺列矩阵 $\widetilde{\boldsymbol{P}}$ 存在，可以表示成

$$\widetilde{\boldsymbol{T}}_5=\boldsymbol{T}_5\widetilde{\boldsymbol{P}} \tag{83}$$

在上例中，

$$\widetilde{\boldsymbol{P}}=[\boldsymbol{e}_2\,\boldsymbol{e}_5\,\boldsymbol{e}_3\,\boldsymbol{e}_1\,\boldsymbol{e}_6\,\boldsymbol{e}_4] \tag{84}$$

而且对于(58)*式的 $\boldsymbol{S}$，作

$$\widetilde{\boldsymbol{S}} \triangleq \widetilde{\boldsymbol{P}}^T\boldsymbol{S}, \tag{85}$$

在上述(a)，(b)中，将“列”换成“行”，将“从左到右”换成“从上到下”，将“$\boldsymbol{b}_j$”换成“$\boldsymbol{h}_j$”后，对于 $\widetilde{\boldsymbol{S}}$ 也同样成立．

这里我们来讨论，用 $\widetilde{\boldsymbol{T}}_5$ 或 $\widetilde{\boldsymbol{T}}_7 \triangleq \widetilde{\boldsymbol{S}}^{-1}$ 进行坐标变换后，矩阵 $\boldsymbol{A}$ 将变成什么样的标准形．在 $\widetilde{\boldsymbol{T}}_5$ 和 $\widetilde{\boldsymbol{T}}_7$ 中，仅是行和列的区别，基本手续完全相似，所以下面我们将介绍一下经常使用的 $\widetilde{\boldsymbol{T}}_7$．

仿照(60)式作 $\widetilde{\boldsymbol{S}}\boldsymbol{A}$．$\widetilde{\boldsymbol{S}}\boldsymbol{A}$ 的第 ν 个子块仅由 $\boldsymbol{h}_i\boldsymbol{A}^{\sigma_i}$ 形的行组成，由(61)式得

$$\boldsymbol{h}_i\boldsymbol{A}^{\sigma_i}=-\sum_{j=1}^{r}\sum_{k=1}^{\sigma_j}\beta_k^{ji}\boldsymbol{e}_{d_j+k}^T\boldsymbol{S}=-\sum_{j=1}^{r}\sum_{k=1}^{\sigma_j}\beta_k^{ji}\boldsymbol{e}_{d_j+k}^T\boldsymbol{P}\boldsymbol{P}^T\boldsymbol{S}. \tag{86}$$

因此，若定义

$$\boldsymbol{e}_{jk}^T \triangleq \boldsymbol{e}_{d_j+k}^T\boldsymbol{P} \tag{87}$$

则

$$\boldsymbol{h}_i\boldsymbol{A}^{\sigma_i}=-\sum_{j=1}^{r}\sum_{k=1}^{\sigma_j}\beta_k^{ji}\boldsymbol{e}_{jk}^T\widetilde{\boldsymbol{S}} \tag{88}$$

故可写成

$$\widetilde{\boldsymbol{S}}\boldsymbol{A} \text{ 的第 } \nu \text{ 个子块}=\boldsymbol{H}\widetilde{\boldsymbol{S}} \tag{89}$$

式中 $\boldsymbol{H}$ 是将 $\beta_k^{ji}(i, j=1, \cdots, r; k=1, \cdots, \sigma_j)$ 适当排列组成的 $r\times n$ 矩阵．

* 原文误为(46)式．——译者注

那么，在整个矩阵的 r 个系列中，有 $(\mu_1-\mu_2)$ 个长度是 1，即 $\boldsymbol{S}$ 中仅含有 $\boldsymbol{h}_j$ 形的行. 在 $\tilde{\boldsymbol{S}}$ 中，根据排列规则 (a)，这些行包含在第 ν 个子块，根据规则 (b)，它们位于第 ν 个子块的最上边. 即，第 ν 个子块整个由 μ_1 行组成，其中从上边起 $(\mu_1-\mu_2)$ 行是 $\boldsymbol{h}_j$ 形的，剩下 μ_2 行是 $\boldsymbol{h}_j\boldsymbol{A}^k(k\geqslant 1)$ 形的. 根据 (a)，(b)，该 μ_2 个行向量，等于对包含在 $\tilde{\boldsymbol{S}}$ 第 $(\nu-1)$ 个子块中的行向量右乘以 $\boldsymbol{A}$ 得到的，即

$$(\tilde{\boldsymbol{S}}\text{ 第 }\nu\text{ 个子块下侧 }\mu_2\text{ 行})=\tilde{\boldsymbol{S}}\text{ 第 }(\nu-1)\text{ 个子块}\times\boldsymbol{A} \tag{90}$$

成立. 整理后得

$$\tilde{\boldsymbol{S}}\boldsymbol{A}\text{ 第 }(\nu-1)\text{ 个子块}=[\underbrace{\;\boldsymbol{0}\;}_{(\mu_1-\mu_2)\text{列}}\;\boldsymbol{I}_{\mu_1}]\times\tilde{\boldsymbol{S}}\text{ 第 }\nu\text{ 个子块} \tag{91}$$

以下根据同样考察可以确证下式

$$\tilde{\boldsymbol{S}}\boldsymbol{A}\text{ 的第 }k\text{ 个子块}=\boldsymbol{U}_k\times\tilde{\boldsymbol{S}}\text{ 的第 }(k+1)\text{ 个子块}\quad k=1,\ \cdots,\ \nu-1 \tag{92}$$

式中 $\boldsymbol{U}_k\triangleq[\boldsymbol{0}\ \boldsymbol{I}]$，其中 $\boldsymbol{0}$ 是 $\mu_{\nu-k+1}\times(\mu_{\nu-k}-\mu_{\nu-k+1})$ 零矩阵，$\boldsymbol{I}$ 是 $\mu_{\nu-k+1}$ 阶单位矩阵.

于是，$\tilde{\boldsymbol{A}}_7\triangleq\tilde{\boldsymbol{S}}\boldsymbol{A}\tilde{\boldsymbol{S}}^{-1}$ 变成下列形式

$$\tilde{\boldsymbol{A}}_7\triangleq\left[\begin{array}{c:cccc}
 & \boldsymbol{U}_1 & & \boldsymbol{0} & \\
\boldsymbol{0} & & \boldsymbol{U}_2 & & \\
 & & & \ddots & \\
 & & & & \boldsymbol{U}_{\nu-1} \\
\hdashline
\multicolumn{5}{c}{\boldsymbol{H}}
\end{array}\right]\begin{array}{l}\}\ \mu_\nu\text{ 行}\\ \}\ \mu_{\nu-1}\text{ 行}\\ \\ \}\ \mu_2\text{ 行}\\ \}\ \mu_1\text{ 行}=r\text{ 行}\end{array} \tag{93}$$

以前关于 σ_j 的选择，只要满足 (i)—(iii)，则是任意的. 若对 σ_j 的选择加以限制，常可得到更简单的标准形.

情况 1　选取 $\sigma_1=\sigma_2=\cdots=\sigma_r$.

这种选择仅在 $\nu=n/r$ 是整数，且

$$\operatorname{rank}[\boldsymbol{B},\ \boldsymbol{A}\boldsymbol{B},\ \cdots,\ \boldsymbol{A}^{\nu-1}\boldsymbol{B}]=n \tag{94}$$

成立的情况下才是可能的. 而且，$\sigma_j=\nu(j=1,\ \cdots,\ r)$.

在这种情况下，因所有系列都具有长度 ν，知

$$\mu_1=\mu_2=\cdots=\mu_\nu=r \tag{95}$$

因此，$\tilde{\boldsymbol{A}}_7$ 中 $\boldsymbol{U}_k$ 均等于 $\boldsymbol{I}_r$. 而且，考虑到(45)式，则

$$\tilde{\boldsymbol{S}}\boldsymbol{B}=\begin{bmatrix}\boldsymbol{0}\\ \cdots\\ \boldsymbol{I}_r\end{bmatrix}\begin{array}{l}\}\ n-r\\ \}\ r\end{array} \tag{96}$$

因此

$$(\boldsymbol{A},\ \boldsymbol{B})\xrightarrow{\tilde{\boldsymbol{T}}_7}(\boldsymbol{A}_9^c,\ \boldsymbol{B}_9^c) \tag{97}$$

$$\boldsymbol{A}_9^c=\left[\begin{array}{c:cccc}
\boldsymbol{0} & \boldsymbol{I}_r & & \boldsymbol{0} & \\
\boldsymbol{0} & & \boldsymbol{I}_r & & \\
 & & & \ddots & \\
 & \boldsymbol{0} & & & \boldsymbol{I}_r \\
\hdashline
\multicolumn{5}{c}{\boldsymbol{H}}
\end{array}\right]=\left[\begin{array}{c:c}
\boldsymbol{0} & \boldsymbol{I}_{n-r}\\
\hdashline
\multicolumn{2}{c}{\boldsymbol{H}}
\end{array}\right],$$

$$B_9^c=\begin{bmatrix}\mathbf{0}\\ \cdots\\ I_r\end{bmatrix}\tag{98}$$

它和单输入情况下的第3典范形相似就更清楚了.

文献[89]中还讲到,任意的 A, B 是否能变换成上面的典范形,一般这是不可能的.

情况 2 $\quad \sigma_j=\text{rank}[b_j,\ Ab_j,\ \cdots,\ A^{n-1}b_j] \quad j=1,\ \cdots,\ r \qquad (99)$

按这个条件选择 σ_j 不一定总是可能的, 因为用(99)式确定出的 σ_j 不一定满足条件(ii), (iii)[1]. 如果可能, 则 $A^{\sigma_j}b_j$ 可仅用 $b_j,\ Ab_j,\ \cdots,\ A^{\sigma_1-1}b_j$ 线性表出. 因此 A_1^c 变成

$$A_{10}^c=\begin{bmatrix}A_{11} & & \mathbf{0}\\ & \ddots & \\ \mathbf{0} & & A_{rr}\end{bmatrix},$$

$$B_{10}^c=\begin{bmatrix}b_{11} & & \mathbf{0}\\ & \ddots & \\ \mathbf{0} & & b_{rr}\end{bmatrix},\tag{100}$$

式中

$$A_{ii}=\left[\begin{array}{cccc|c}0 & & & \mathbf{0} & \alpha_1^{ii}\\ 1 & \ddots & & & \vdots\\ & \ddots & \ddots & 0 & \alpha_{\sigma_i-1}^{ii}\\ \mathbf{0} & & & 1 & \alpha_{\sigma_i}^{ii}\end{array}\right]=\sigma_i\times\sigma_i,$$

$$b_{ii}=\begin{bmatrix}1\\0\\ \vdots\\ \\0\end{bmatrix}=\sigma_i\times 1\tag{101}$$

它归结到相互独立的 r 个单输入系统的集合. 以下,对各单输入系统用前节的变换即可. (100), (101)式结构的标准形,若存在则唯一确定.

情况 3 $\quad$ 设 $B_k\triangleq[b_1,\ b_2,\ \cdots,\ b_k] \quad 1\leqslant k\leqslant r-1;\ B_r\triangleq B \qquad (102)$

σ_j 按如下确定

$$\sigma_1\triangleq\text{rank}[b_1,\ Ab_1,\ \cdots,\ A^{n-1}b_1]\tag{103}$$

$$\begin{aligned}\sigma_j\triangleq&\text{rank}[B_j,\ AB_j,\ \cdots,\ A^{n-1}B_j]\\ &-\text{rank}[B_{j-1},\ AB_{j-1},\ \cdots,\ A^{n-1}B_{j-1}] \quad 2\leqslant j\leqslant r\end{aligned}\tag{104}$$

由定义立即得

$$d_i\triangleq\sum_{j=1}^{i}\sigma_j=\text{rank}[B_i,\ AB_i,\ \cdots,\ A^{n-1}B_i]\tag{105}$$

由此显而易见,若可控,则按上面选择的 σ_j 能够满足条件(ii), (iii). 而

$$\begin{aligned}&\text{rank}[B_j,\ AB_j,\ \cdots,\ A^{n-1}B_j]\\ &\quad>\text{rank}[B_{j-1},\ AB_{j-1},\ \cdots,\ A^{n-1}B_{j-1}]\ \forall j\end{aligned}\tag{106}$$

---

1) 从第II部分将要讲的几何学的角度来看, 当设 $(A,\ b_j)$ 可控的部分空间为 $S_j(j=1,\ \cdots,\ r)$ 时, 状态空间可以分解成 S_j 的直和,是用(99)式确定的 σ_j 满足(i)～(iii)的充分必要条件.

是剩下的条件(i)成立的充分必要条件．若上式不成立，即若假定对于某个 $2\leqslant k\leqslant r$

$$\begin{aligned}&\operatorname{rank}[\boldsymbol{B}_k,\ \boldsymbol{AB}_k,\ \cdots,\ \boldsymbol{A}^{n-1}\boldsymbol{B}_k]\\&=\operatorname{rank}[\boldsymbol{B}_{k-1},\ \boldsymbol{AB}_{k-1},\ \cdots,\ \boldsymbol{A}^{n-1}\boldsymbol{B}_{k-1}],\end{aligned}\tag{107}$$

则由 $\boldsymbol{B}$ 中去掉第 k 列 $\boldsymbol{b}_k$，即去掉第 k 个输入 u_k，系统仍然可控．因此，系统在利用全部输入时可控，而去掉其中任意一个输入都会成为不可控的，这就成为条件(i)成立的充分条件．以下为了简单起见，假定满足这个条件．

根据 σ_j 的选择方法，在这种情况下 $\boldsymbol{A}^{\sigma_j}\boldsymbol{b}_j$ 可以用向量

$$\boldsymbol{b}_1,\ \boldsymbol{Ab}_1,\ \cdots,\ \boldsymbol{A}^{\sigma_1-1}\boldsymbol{b}_1;\ \cdots;\ \boldsymbol{b}_j,\ \boldsymbol{Ab}_j,\ \cdots,\ \boldsymbol{A}^{\sigma_j-1}\boldsymbol{b}_j\tag{108}$$

线性表出，所以

$$\alpha_k^{ji}=0\quad j>i;\ k=1,\ \cdots,\ \sigma_j\tag{109}$$

成立．因此，在 $\boldsymbol{A}_5^c$ 中 $\boldsymbol{A}_{ji}(j>i)$ 均为 $\mathbf{0}$，$\boldsymbol{A}_5^c$ 变成分块上三角形矩阵．

$$\boldsymbol{A}_1^c=\begin{bmatrix}\boldsymbol{A}_{11}&\boldsymbol{A}_{12}&\cdots\cdots&\boldsymbol{A}_{1r}\\&\boldsymbol{A}_{22}&\cdots\cdots&\boldsymbol{A}_{2r}\\&&\ddots&\vdots\\\mathbf{0}&&&\boldsymbol{A}_{rr}\end{bmatrix}\quad \boldsymbol{B}_{11}^c=\boldsymbol{B}_5^c\qquad \boldsymbol{A}_{ii},\ \boldsymbol{A}_{ij}\ \text{分别为(50)式形式的矩阵}\tag{110}$$

容易看到，在第 2 典范形 $\boldsymbol{A}_6^c$ 中，也是 $\boldsymbol{A}'_{ij}=\mathbf{0}(i>j)$，即

$$\boldsymbol{A}_6^c=\begin{bmatrix}\boldsymbol{A}'_{11}&\boldsymbol{A}'_{12}&\cdots\cdots&\boldsymbol{A}'_{1r}\\&\boldsymbol{A}'_{22}&&\vdots\\&&\ddots&\vdots\\\mathbf{0}&&&\boldsymbol{A}'_{rr}\end{bmatrix}\tag{111}$$

所以，一般 $(\boldsymbol{A}_6^c)^k\ k\geqslant2$ 也变成同样形式的分块上三角形矩阵．因此

$$\boldsymbol{e}_{d_i+1}^T(\boldsymbol{A}_6^c)^k=[\underbrace{0\cdots0}_{d_i}\ \underbrace{*\cdots*}_{n-d_i}]\quad k=0,\ 1,\ \cdots,\ \sigma_i-1;\ i=1,\ \cdots,\ r\tag{112}$$

将上式代入 $\boldsymbol{R}$，可见 $\boldsymbol{R}$ 也是形状和 $\boldsymbol{A}_6^c$ 相同的分块上三角形矩阵．因此，在 $\boldsymbol{A}_7^c=\boldsymbol{RA}_6^c\boldsymbol{R}^{-1}$ 中，右边三个矩阵都是同样形状的分块上三角形矩阵，其积当然是同样形状的分块上三角形矩阵．于是可以看到，在 $\boldsymbol{A}_7^c$ 中也是 $\boldsymbol{A}''_{ij}=\mathbf{0}(i>j)$．而且，若将(55)式的 $\boldsymbol{B}_6^c$ 和(112)式代入 $\boldsymbol{B}_7^c=\boldsymbol{RB}_6^c$，可以看到这种情况下 $\boldsymbol{B}_7^c$ 是

$$\boldsymbol{B}_7^c=\begin{bmatrix}\boldsymbol{B}''_1\\\cdots\\\boldsymbol{B}''_2\\\cdots\\\vdots\\\cdots\\\boldsymbol{B}''_r\end{bmatrix}\begin{matrix}\}\sigma_1\\\\\}\sigma_2\\\\\\\\\}\sigma_r\end{matrix}\qquad \boldsymbol{B}''_i=[\underbrace{\ \mathbf{0}\ }_{(i-1)\text{列}}\ \vdots\ *]=\sigma_i\times r\quad i=1,\ \cdots,\ r\tag{113}$$

形式的矩阵．在 $\boldsymbol{A}_8^c$ 中直接可以看出，δ_l^{ji} 用(75)式给出，$\alpha_k^{ji}=0(j>i,\ ^\forall k)\Rightarrow\delta_l^{ji}=0(j>i,\ ^\forall l)$ 成立．因此，$\tilde{\boldsymbol{A}}_{ji}=\mathbf{0}(j>i)$，$\boldsymbol{A}_8^c$ 也成为分块上三角形矩阵[90]．

情况 4[89]　在可控系统中，矩阵 $\boldsymbol{Q}_n \triangleq [\boldsymbol{B}, \boldsymbol{AB}, \cdots, \boldsymbol{A}^{n-1}\boldsymbol{B}]$ 包含有 n 个线性独立的列. 在这里，我们按如下规则从 $\boldsymbol{Q}_n$ 中选取 n 个线性独立的列. 亦即假定，当 $\boldsymbol{Q}_n$ 的第 j 列不能用位于其左边的列(即第 1 列，第 2 列，……，第 $(j-1)$ 列)线性表出时则采用；若能用位于其左边的列线性表出时则不采用. 在这种情况下，若 $\boldsymbol{Q}_n$ 的列 $\boldsymbol{A}^k\boldsymbol{b}_j(1\leqslant j\leqslant r)$ 可以用位于其左边的列线性表出，则显然 $\boldsymbol{A}^l\boldsymbol{b}_j(l\geqslant k)$ 也可以用位于其左边的列线性表出. 即，若 $\boldsymbol{A}^k\boldsymbol{b}_j$ 不采用，则 $\boldsymbol{A}^l\boldsymbol{b}_j(l\geqslant k)$ 亦不采用. 因此，将所采用的列的顺序改换后，可以得到下列形式的矩阵

$$\begin{aligned}\boldsymbol{T}_{12}=[&\boldsymbol{b}_1,\ \boldsymbol{Ab}_1,\ \cdots,\ \boldsymbol{A}^{\sigma_1-1}\boldsymbol{b}_1 \mid \boldsymbol{b}_2,\ \boldsymbol{Ab}_2,\ \cdots,\ \boldsymbol{A}^{\sigma_2-1}\boldsymbol{b}_2 \mid \\ &\cdots \mid \boldsymbol{b}_r,\ \boldsymbol{Ab}_r,\ \cdots,\ \boldsymbol{A}^{\sigma_r-1}\boldsymbol{b}_r]\end{aligned} \tag{114}$$

显然，若 $\operatorname{rank}\boldsymbol{B}=r$ 是可控的，则 $\sigma_1,\ \cdots,\ \sigma_r$ 及 $\boldsymbol{T}_{12}$ 满足(i)，(ii)，(iii).

而且，当设 α 为前一章所定义的可控性指标时，则有 $\alpha=\max\limits_j \sigma_j$ 的关系.

那么在这种情况下，$\boldsymbol{A}^k\boldsymbol{b}_j(k\geqslant\sigma_j)$ 可以用下列向量线性表出：

$$\begin{aligned}&\boldsymbol{b}_1,\ \boldsymbol{Ab}_1,\ \cdots,\ \boldsymbol{A}^{\rho_1-1}\boldsymbol{b}_1;\ \boldsymbol{b}_2,\ \boldsymbol{Ab}_2,\ \cdots,\ \boldsymbol{A}^{\rho_2-1}\boldsymbol{b}_2;\ \cdots;\\ &\boldsymbol{b}_r,\ \boldsymbol{Ab}_r,\ \cdots,\ \boldsymbol{A}^{\rho_r-1}\boldsymbol{b}_r\end{aligned} \tag{115}$$

式中

$$\rho_l \triangleq \begin{cases}\min\{\sigma_l,\ k+1\} & 1\leqslant l<j\\ \min\{\sigma_l,\ k\} & j\leqslant l\leqslant r.\end{cases} \tag{116}$$

这一点对于典范形 $(\boldsymbol{A}_7^c,\ \boldsymbol{B}_7^c)$ 特别重要.

现在，再稍微讲一下(57)式中的 $*$. 因 $\boldsymbol{A}^k\boldsymbol{b}_j(k\geqslant\sigma_j)$ 可用上述向量线性表出，则

$$\boldsymbol{h}_i\boldsymbol{A}^k\boldsymbol{b}_j=\boldsymbol{h}_i\sum_{q=1}^{r}\sum_{p=0}^{\rho_q-1}C_{p,q}\boldsymbol{A}^p\boldsymbol{b}_q=\sum_{q=1}^{r}\sum_{p=0}^{\rho_q-1}C_{p,q}\boldsymbol{h}_i\boldsymbol{A}^p\boldsymbol{b}_q \tag{117}$$

在这里，若用(57)式的第 1，2 式，则得

$$\boldsymbol{h}_i\boldsymbol{A}^p\boldsymbol{b}_q=0\quad p=0,\ \cdots,\ \rho_q-1;\ q=1,\ \cdots,\ r;\ q\neq i \tag{118}$$

而且

$$\boldsymbol{h}_i\boldsymbol{A}^p\boldsymbol{b}_i=0\quad p=0,\ \cdots,\ \rho_i-2 \tag{119}$$

$$\boldsymbol{h}_i\boldsymbol{A}^{\rho_i-1}\boldsymbol{b}_i=\begin{cases}1 & \rho_i=\sigma_i,\ \text{即当 } 1\leqslant i<j \text{ 时},\ \sigma_i\leqslant k+1\\ & \text{或当 } j<i\leqslant r \text{ 时},\ \sigma_i\leqslant k\\ 0 & \rho_i<\sigma_i,\ \text{即当 } 1\leqslant i<j \text{ 时},\ k<\sigma_i-1\\ & \text{或当 } j<i\leqslant r \text{ 时},\ k<\sigma_i\end{cases} \tag{120}$$

因此，有比(57)式还强的如下性质成立：

$$\boldsymbol{h}_i\boldsymbol{A}^k\boldsymbol{b}_j=\begin{cases}0 & 0\leqslant k<\sigma_j-1\\ 0 & j\neq i\quad k=\sigma_j-1\\ 0 & k<\sigma_i-1\\ 1 & j=i\quad k=\sigma_i-1\\ * & j\neq i\quad k\geqslant\max\{\sigma_j,\ \sigma_i-1\}\end{cases} \tag{121}$$

我们现在利用该性质来计算 $\boldsymbol{SB}$. $\boldsymbol{SB}$ 的第 1 行—第 (σ_1-1) 行具有

$$\boldsymbol{h}_1\boldsymbol{A}^k\boldsymbol{B}\quad k=0,\ 1,\ \cdots,\ \sigma_1-2 \tag{122}$$

的形式，而由(121)式的第 3 式得

$$\boldsymbol{h}_1\boldsymbol{A}^k\boldsymbol{B}=0\quad k=0,\ 1,\ \cdots,\ \sigma_1-2. \tag{123}$$

其次，第 σ_1 行是 $\boldsymbol{h}_1\boldsymbol{A}^{\sigma_1-1}\boldsymbol{B}$. 由 (121) 式的第 4 式，这里 $\boldsymbol{h}_1\boldsymbol{A}^{\sigma_1-1}\boldsymbol{b}_1$ 等于 1. 若 $\sigma_1-1\leqslant\sigma_j-1$, 即 $\sigma_1\leqslant\sigma_j$, 由第 1, 2 式，则 $\boldsymbol{h}_1\boldsymbol{A}^{\sigma_1-1}\boldsymbol{b}_j(j=2,\ \cdots,\ r)$ 是 0. 若 $\sigma_j<\sigma_1$, 由第 5 式，则一般成为不一定为 0 的某个数. 若继续论证下去，则最后 $\boldsymbol{B}_7^c$ 变成如下形式的 $\boldsymbol{B}_{12}^c$.

$$\boldsymbol{B}_{12}^c=\boldsymbol{S}\boldsymbol{B}=\begin{bmatrix}\boldsymbol{B}_1\\ \cdots\\ \boldsymbol{B}_2\\ \cdots\\ \vdots\\ \cdots\\ \boldsymbol{B}_r\end{bmatrix},\qquad \boldsymbol{B}_i\triangleq\begin{bmatrix}\boldsymbol{0}\\ \cdots\cdots\cdots\cdots\\ b_{i1},\ \cdots,\ b_{ir}\end{bmatrix}=\sigma_i\times r\quad i=1,\ \cdots,\ r$$

$$b_{ij}=\begin{cases}0 & j\neq i\quad \sigma_j\geqslant\sigma_i\\ 1 & j=i\\ * & j\neq i\quad \sigma_j<\sigma_i\end{cases}\tag{124}$$

在这里，若定义

$$\boldsymbol{B}^*=\begin{bmatrix}b_{11}\cdots b_{1r}\\ \vdots\qquad\vdots\\ b_{r1}\cdots b_{rr}\end{bmatrix}=r\times r\tag{125}$$

按照如下证明，$\boldsymbol{B}^*$ 是正则的.

$$\operatorname{rank}\boldsymbol{B}^*=\operatorname{rank}\boldsymbol{B}_{12}^c=\operatorname{rank}\boldsymbol{B}=r\tag{126}$$

其次，设 $\bar{b}_{ii}\triangleq 0$,

$$\text{若 } b_{ij}=0,\ \text{则 } \bar{b}_{ij}\triangleq 0\quad i\neq j\tag{127}$$

$$\text{若 } b_{ij}=*,\ \text{则 } \bar{b}_{ij}\triangleq 1\quad i\neq j\tag{128}$$

作布尔矩阵 $\overline{\boldsymbol{B}}^*\triangleq[\bar{b}_{ij}]$. 结合矩阵为 $\overline{\boldsymbol{B}}^*$ 的有向图中没有回路存在. 其原因是，若有回路存在，则对于某个 m 和 $j_1,\ j_2,\ \cdots,\ j_m$, 得

$$\bar{b}_{j_1j_2}\bar{b}_{j_2j_3}\cdots\bar{b}_{j_{m-1}j_m}\bar{b}_{j_mj_1}=1\tag{129}$$

因此

$$\bar{b}_{j_1j_2}=\bar{b}_{j_2j_3}=\cdots=\bar{b}_{j_mj_1}=1,\tag{130}$$

即

$$\sigma_{j_1}>\sigma_{j_2}>\cdots>\sigma_{j_{m-1}}>\sigma_{j_m}>\sigma_{j_1},\tag{131}$$

产生矛盾. 因为不具有回路，利用适当的顺列矩阵 $\boldsymbol{P}_0$, 则可以将 $\boldsymbol{P}_0^T\boldsymbol{B}^*\boldsymbol{P}_0$ 变换成主对角线上的元素为 1 的上三角形矩阵.

$\boldsymbol{A}_{12}^c$ 是和 $\boldsymbol{A}_7^c$ 具有相同结构的矩阵.

容易证明，不用 $\boldsymbol{S}$ 而用 $\widetilde{\boldsymbol{S}}$, 若进行同样的考察，设 $\boldsymbol{T}\triangleq\widetilde{\boldsymbol{S}}^{-1}$, 则

$$(\boldsymbol{A},\ \boldsymbol{B})\xrightarrow{\boldsymbol{T}}(\boldsymbol{A}_{13}^c,\ \boldsymbol{B}_{13}^c)\tag{132}$$

$$\boldsymbol{A}_{13}^c=\left[\begin{array}{c:cccc}\boldsymbol{0} & \boldsymbol{U}_1 & & & \boldsymbol{0}\\ & & \boldsymbol{U}_2 & & \\ & & & \ddots & \\ & \boldsymbol{0} & & & \boldsymbol{U}_{\nu-1}\\ \hdashline \multicolumn{5}{c}{\boldsymbol{H}}\end{array}\right],\qquad \boldsymbol{B}_{13}^c=\begin{bmatrix}\boldsymbol{0}\\ \cdots\cdots\\ \boldsymbol{B}^*\end{bmatrix}\tag{133}$$

到现在为止,我们讨论的都是状态空间的座标变换,在这上面也可以再加上输入空间的正则座标变换,即

$$\boldsymbol{\mu}=\boldsymbol{G}\bar{\boldsymbol{u}} \quad \boldsymbol{G}\in R^{r\times r} \quad \det \boldsymbol{G}\neq 0. \tag{134}$$

若将这样的变换用$(\boldsymbol{T}, \boldsymbol{G})$表示,则

$$(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D})\xrightarrow{(\boldsymbol{T}, \boldsymbol{G})}(\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}, \boldsymbol{T}^{-1}\boldsymbol{B}\boldsymbol{G}, \boldsymbol{C}\boldsymbol{T}, \boldsymbol{D}\boldsymbol{G}) \tag{135}$$

对于变换$(\boldsymbol{T}, \boldsymbol{G})$,系统的可控性、可观测性,$\boldsymbol{A}$ 的特征值是不变的,但是输入输出关系已经不是不变了.

引入变换 $\boldsymbol{G}$ 的效果是,无论在$(\boldsymbol{A}_{12}^{c}, \boldsymbol{B}_{12}^{c})$或$(\boldsymbol{A}_{13}^{c}, \boldsymbol{B}_{13}^{c})$中,若设 $\boldsymbol{G}_0\triangleq\boldsymbol{B}^{*-1}$,则均可将 $\boldsymbol{B}$ 变换成非常简单的形状.

$$(\boldsymbol{A}, \boldsymbol{B})\xrightarrow{(\boldsymbol{T}_{12}, \boldsymbol{G}_0)}(\boldsymbol{A}_{14}^{c}, \boldsymbol{B}_{14}^{c}) \tag{136}$$

$$(\boldsymbol{A}, \boldsymbol{B})\xrightarrow{(\boldsymbol{T}_{13}, \boldsymbol{G}_0)}(\boldsymbol{A}_{15}^{c}, \boldsymbol{B}_{15}^{c}) \tag{137}$$

式中 $\boldsymbol{A}_{14}^{c}$ 和 $\boldsymbol{A}_{7}^{c}$ 具有相同结构,$\boldsymbol{B}_{14}^{c}$ 和 $\boldsymbol{B}_{12}^{c}$ 具有相同结构,特别是

$$\boldsymbol{B}_i=\begin{bmatrix}\boldsymbol{0}\\ \boldsymbol{e}_i^T\end{bmatrix} \tag{138}$$

而且,$\boldsymbol{A}_{15}^{c}$ 和 $\boldsymbol{A}_{13}^{c}$ 具有相同结构,$\boldsymbol{B}_{15}^{c}$ 为

$$\boldsymbol{B}_{15}^{c}=\begin{bmatrix}\boldsymbol{0}\\ \cdots\\ \boldsymbol{I}_r\end{bmatrix}. \tag{139}$$

在文献[92]中用过这个典范形.

其次,我们来看一下 $\boldsymbol{A}_{11}^{c}$. 首先,进行$(\boldsymbol{I}, \boldsymbol{G})$变换,变成 $\boldsymbol{B}\longrightarrow\boldsymbol{B}\boldsymbol{G}$. 而后假定,对 $\boldsymbol{A}$, $\boldsymbol{B}\boldsymbol{G}$ 适用(103)—(105)式的规则,作变换矩阵 $\boldsymbol{T}_G$. 即假定 $\boldsymbol{G}$ 的第 j 列为 $\boldsymbol{g}_j$,$\boldsymbol{G}_k\triangleq[\boldsymbol{g}_1, \cdots, \boldsymbol{g}_k]$ $1\leqslant k\leqslant\nu-1$,$\boldsymbol{G}_\nu\triangleq\boldsymbol{G}$,

$$\bar{\sigma}_1\triangleq \operatorname{rank}\,[\boldsymbol{B}\boldsymbol{g}_1, \boldsymbol{A}\boldsymbol{B}\boldsymbol{g}_1, \cdots, \boldsymbol{A}^{n-1}\boldsymbol{B}\boldsymbol{g}_1] \tag{140}$$

$$\begin{aligned}\bar{\sigma}_j\triangleq& \operatorname{rank}\,[\boldsymbol{B}\boldsymbol{G}_j, \boldsymbol{A}\boldsymbol{B}\boldsymbol{G}_j, \cdots, \boldsymbol{A}^{n-1}\boldsymbol{B}\boldsymbol{G}_j]\\ &-\operatorname{rank}[\boldsymbol{B}\boldsymbol{G}_{j-1}, \boldsymbol{A}\boldsymbol{B}\boldsymbol{G}_{j-1}, \cdots, \boldsymbol{A}^{n-1}\boldsymbol{B}\boldsymbol{G}_{j-1}].\\ &2\leqslant j\leqslant r\end{aligned} \tag{141}$$

一般,若 $\boldsymbol{G}$ 不同,则 $\bar{\sigma}_j$ 的值也不同. 因此,在用

$$(\boldsymbol{A}, \boldsymbol{B})\xrightarrow{(\boldsymbol{T}_G, \boldsymbol{G})}(\boldsymbol{A}_{16}^{c}, \boldsymbol{B}_{16}^{c}) \tag{142}$$

得到的典范形中,虽然 $\boldsymbol{A}_{16}^{c}$ 具有 $\boldsymbol{A}_{11}^{c}$ 的结构,但是一般子矩阵的阶数及非零元素的值不同. 于是可以看到,具有 $\boldsymbol{A}_{11}^{c}$ 的构造的典范形,在变换 $\boldsymbol{T}$ 的范围内唯一确定,但是将变换范围扩大到$(\boldsymbol{T}, \boldsymbol{G})$就已经不是唯一了.

情况 1, 2 的选择,$\boldsymbol{A}$, $\boldsymbol{B}$ 必须满足一定的条件. 但是,将容许变换的范围扩大到$(\boldsymbol{T}, \boldsymbol{G})$情况又如何呢? 首先,对于情况 1,既使扩大变换范围,n, r 也不变化. 而且,对于 $\nu=n/r$,因

$$\operatorname{rank}[\boldsymbol{B}\boldsymbol{G}, \boldsymbol{A}\boldsymbol{B}\boldsymbol{G}, \cdots, \boldsymbol{A}^{\nu-1}\boldsymbol{B}\boldsymbol{G}]=\operatorname{rank}[\boldsymbol{B}, \boldsymbol{A}\boldsymbol{B}, \cdots, \boldsymbol{A}^{\nu-1}\boldsymbol{B}] \tag{143}$$

则当且仅当在变换 $\boldsymbol{T}$ 的范围内情况 1 的变换可能时,在$(\boldsymbol{T}, \boldsymbol{G})$的范围内是可能的. 而且,标准形、阶数是唯一的.

情况 2 既使在变换 T 的范围内不可能，但是扩大到 (T, G) 后常是可能的. 例如

$$A \triangleq \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad B \triangleq \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \tag{144}$$

变成

$$\operatorname{rank}[b_1, \ Ab_1, \ A^2 b_1] = \operatorname{rank}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} = 3,$$

$$\operatorname{rank}[b_2, \ Ab_2, \ A^2 b_2] = \operatorname{rank}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} = 1, \ 3+1=4>n,$$

情况 2 的变换不可能. 但是，若设 $G \triangleq \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$，则

$$(A, \ B) \xrightarrow{(T, \ G)} (A_{10}^c, \ B_{10}^c)$$

是可能的.

A-I-13 状态反馈

对于常系数线性系统

$$\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{u},\ \boldsymbol{y}=\boldsymbol{C}\boldsymbol{x};\ \boldsymbol{x}(t)\in R^n,\ \boldsymbol{u}(t)\in R^r,\ \boldsymbol{y}(t)\in R^m,$$
$$\boldsymbol{A}\in R^{n\times n},\ \ \boldsymbol{B}\in R^{n\times r},\ \ \boldsymbol{C}\in R^{m\times n} \tag{1}$$

控制输入 $\boldsymbol{u}$ 试用

$$\boldsymbol{u}=\boldsymbol{K}\boldsymbol{x}+\boldsymbol{v},\ \boldsymbol{v}(t)\in R^r,\ \boldsymbol{K}\in R^{r\times n} \tag{2}$$

给出. 由图 1 也可以看出，将(1)，(2)式合起来，表示在(1)系统上加上状态变量 $\boldsymbol{x}$ 的线性组合反馈形成的"闭环系统"的特性. 因此，将(2)式称为状态反馈(state feedback). 在同样意义上，称下式

$$\boldsymbol{u}=\boldsymbol{H}\boldsymbol{y}+\boldsymbol{v},\ \boldsymbol{v}(t)\in R^r,\ \boldsymbol{H}\in R^{r\times m} \tag{3}$$

为输出反馈(output feedback).

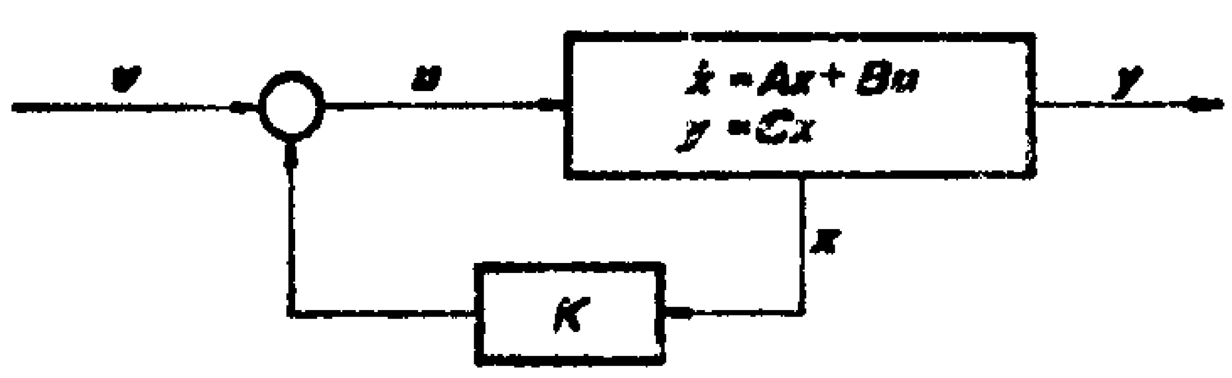


图 1

以下，我们把这种状态反馈在控制系统上所起的各种作用，作为说明以上讲过的矩阵诸性质应用的例子. 但是，大家对于图 1 可能产生下列疑问. 即，$\boldsymbol{x}$ 是(1)式系统的内部状态，一般不能直接观测，而可以观测的只有输出 $\boldsymbol{y}$(参看 A-I-1). 那么，由 $\boldsymbol{x}$, $\boldsymbol{v}$ 确定 $\boldsymbol{u}$ 的控制规律((2)式)是不是就没有实用意义了. 关于这一点，为了由可观测的 $\boldsymbol{y}$ 估计 $\boldsymbol{x}$ 值，建议采用观测器(observer). 使用它可以使上述问题在一定程度上得到解决. 下一章将介绍这种观测器的理论.

反馈变换

将(2)式代入(1)式，得

$$\dot{\boldsymbol{x}}=(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K})\boldsymbol{x}+\boldsymbol{B}\boldsymbol{v},\ \boldsymbol{y}=\boldsymbol{C}\boldsymbol{x} \tag{4}$$

该式和(1)式具有同样形式. 因此，(2)式的状态反馈也可以看成是矩阵组($\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$)到($\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K}$, $\boldsymbol{B}$, $\boldsymbol{C}$)的一种变换，通常称其为反馈变换. 我们将它和已经多次使用的状态空间的坐标变换、输入空间的坐标变换结合起来，定义变换($\boldsymbol{T}$, $\boldsymbol{G}$, $\boldsymbol{K}$)

$$(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{C})\xrightarrow{(\boldsymbol{T},\ \boldsymbol{G},\ \boldsymbol{K})}(\boldsymbol{T}^{-1}(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K})\boldsymbol{T},\ \boldsymbol{T}^{-1}\boldsymbol{B}\boldsymbol{G},\ \boldsymbol{C}\boldsymbol{T}) \tag{5}$$

由(5)式可见，若包含反馈变换，则必须考虑变换的顺序. 这里变换($\boldsymbol{T}$, $\boldsymbol{G}$, $\boldsymbol{K}$)定义成，按照反馈变换 $\boldsymbol{K}$, 状态变换 $\boldsymbol{T}$, 输入变换 $\boldsymbol{G}$ 的顺序进行.

(i) 先进行变换($\boldsymbol{T}_1$, $\boldsymbol{G}_1$, $\boldsymbol{K}_1$)，然后再进行变换($\boldsymbol{T}_2$, $\boldsymbol{G}_2$, $\boldsymbol{K}_2$)，与进行($\boldsymbol{T}_1\boldsymbol{T}_2$, $\boldsymbol{G}_1\boldsymbol{G}_2$, $\boldsymbol{K}_1+\boldsymbol{G}_1\boldsymbol{K}_2\boldsymbol{T}_1^{-1}$)变换等价. 即

$$(\boldsymbol{T}_2,\ \boldsymbol{G}_2,\ \boldsymbol{K}_2)\circ(\boldsymbol{T}_1,\ \boldsymbol{G}_1,\ \boldsymbol{K}_1)=(\boldsymbol{T}_1\boldsymbol{T}_2,\ \boldsymbol{G}_1\boldsymbol{G}_2,\ \boldsymbol{K}_1+\boldsymbol{G}_1\boldsymbol{K}_2\boldsymbol{T}_1^{-1}) \tag{6}$$

(ii) 设变换$(\boldsymbol{T},\ \boldsymbol{G},\ \boldsymbol{K})$全体的集合为$G$(其中$\boldsymbol{T}=n\times n$, $\boldsymbol{G}=r\times r$, $\boldsymbol{K}=r\times n$, $\boldsymbol{T}$, $\boldsymbol{G}$是正则的)，在G的元之间用(6)式定义“积”$\circ$，则G是关于$\circ$的群.

(iii) 当两个系统$(\boldsymbol{A}_1, \boldsymbol{B}_1, \boldsymbol{C}_1)$和$(\boldsymbol{A}_2, \boldsymbol{B}_2, \boldsymbol{C}_2)$可以用$G$中某个元表示成

$$(\boldsymbol{A}_1,\ \boldsymbol{B}_1,\ \boldsymbol{C}_1)\xrightarrow{(\boldsymbol{T},\ \boldsymbol{G},\ \boldsymbol{K})}(\boldsymbol{A}_2,\ \boldsymbol{B}_2,\ \boldsymbol{C}_2) \tag{7}$$

时，称为两个系统是G等价的. G等价是等值关系.

(iv) 利用(6)式的“积”$\circ$，可以将G中任意元$(\boldsymbol{T},\ \boldsymbol{G},\ \boldsymbol{K})$分解成

$$(\boldsymbol{T},\ \boldsymbol{G},\ \boldsymbol{K})=(\boldsymbol{T},\ \boldsymbol{I}_r,\ \boldsymbol{0})\circ(\boldsymbol{I}_n,\ \boldsymbol{G},\ \boldsymbol{0})\circ(\boldsymbol{I}_n,\ \boldsymbol{I}_r,\ \boldsymbol{K}) \tag{8}$$

可控性和可观测性

我们已经知道下面感兴趣的结果.

[定理 1] 对于G等价的系统，可控性保存.

为了证明该定理，只要证明对于(8)式右边各项可控性保存即可. 因为对于$(\boldsymbol{I}_n,\ \boldsymbol{G},\ \boldsymbol{0})$, $(\boldsymbol{T},\ \boldsymbol{I}_r,\ \boldsymbol{0})$是显然的，故只证明对于$(\boldsymbol{I}_n,\ \boldsymbol{I}_r,\ \boldsymbol{K})$，即对于反馈变换可控性是保存的. 在这里介绍两个证明方法.

(证明方法 1) 设(1)式系统完全可控，当起始状态固定在任意值$\boldsymbol{x}_0$时，则有在某个$t_f>0$使$\boldsymbol{x}(t_f)=\boldsymbol{0}$的控制作用$\boldsymbol{u}^*_{[0,t_f]}$存在. $\boldsymbol{u}^*_{[0,t_f]}$使状态循着$\boldsymbol{x}^*_{[0,t_f]}$到达原点. 因此，若

$$\boldsymbol{v}^*(t)\triangleq\boldsymbol{u}^*(t)-\boldsymbol{K}\boldsymbol{x}^*(t)\quad 0\leqslant t\leqslant t_f, \tag{9}$$

则在以$\boldsymbol{x}_0$为起始状态的“闭环系统”((4)式)上，将$\boldsymbol{v}^*_{[0,t_f]}$作为输入加上时，显然$\boldsymbol{x}(t_f)=\boldsymbol{0}$. 即，当$(\boldsymbol{A},\ \boldsymbol{B})$完全可控时，$(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K},\ \boldsymbol{B})$也完全可控[1]. 反之也可以同样证明.

证明完毕

(证明方法 2) $(\boldsymbol{A},\ \boldsymbol{B})$的可控性矩阵(参看 A-I-11)是

$$\boldsymbol{Q}_n=[\boldsymbol{B}\,\vdots\,\boldsymbol{A}\boldsymbol{B}\,\vdots\,\cdots\,\vdots\,\boldsymbol{A}^{n-1}\boldsymbol{B}], \tag{10}$$

$(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K},\ \boldsymbol{B})$的可控性矩阵是

$$\overline{\boldsymbol{Q}}_n=[\boldsymbol{B}\,\vdots\,(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K})\boldsymbol{B}\,\vdots\,\cdots\,\vdots\,(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K})^{n-1}\boldsymbol{B}]. \tag{11}$$

$\overline{\boldsymbol{Q}}_n$的第 1 个子块等于$\boldsymbol{Q}_n$的第 1 个子块，$\overline{\boldsymbol{Q}}_n$的第 2 个子块若展开成

$$(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K})\boldsymbol{B}=\boldsymbol{A}\boldsymbol{B}+\boldsymbol{B}(\boldsymbol{K}\boldsymbol{B}), \tag{12}$$

则可以看到，它是在$\boldsymbol{Q}_n$第 2 个子块$(\boldsymbol{A}\boldsymbol{B})$上加上$\boldsymbol{Q}_n$第 1 个子块的常数倍. 若$\overline{\boldsymbol{Q}}_n$的第 3 个子块也展开成

$$(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K})^2\boldsymbol{B}=\boldsymbol{A}^2\boldsymbol{B}+\boldsymbol{A}\boldsymbol{B}(\boldsymbol{K}\boldsymbol{B})+\boldsymbol{B}(\boldsymbol{K}\boldsymbol{A}\boldsymbol{B})+\boldsymbol{B}(\boldsymbol{K}\boldsymbol{B}\boldsymbol{K}\boldsymbol{B}), \tag{13}$$

则可以看到，它是在$\boldsymbol{Q}_n$第 3 个子块$(\boldsymbol{A}^2\boldsymbol{B})$上加上$\boldsymbol{Q}_n$第 1，第 2 个子块的常数倍. 以下同样. 因将矩阵某列的常数倍加到其它列后矩阵的秩不变(参看 B-I-3)，则

$$\operatorname{rank}\boldsymbol{Q}_n=\operatorname{rank}\overline{\boldsymbol{Q}}_n, \tag{14}$$

因而证明

$$(\boldsymbol{A},\ \boldsymbol{B})\text{可控}\Leftrightarrow(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K},\ \boldsymbol{B})\text{可控}.$$

证明完毕

---

1) 在 A-I-12 中，为了表示系统(1)引入符号$(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{C})$. 同样，当不考虑输出如何，而只关心状态方程式时，代替“系统$\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{u}$”的写法，用符号$(\boldsymbol{A},\ \boldsymbol{B})$表示.

［定理 2］ 对于输出反馈，可控性、可观测性保存.

（证明） 因输出反馈不过是状态反馈的特殊情况（$\boldsymbol{K}=\boldsymbol{HC}$），由定理 1，显然可控性保存.

对于可观测性，因(3)式中输出 $\boldsymbol{y}$，输入 $\boldsymbol{v}$ 皆由观测得知，则 $\boldsymbol{u}$ 是“已知输入”．因已知输入的存在对可观测性无任何影响(参看 A-I-8)，则可观测性保存.

证明完毕

（问题 1） 为何对于状态反馈可观测性不一定保存？

可控系统的典范形

前章讨论过，利用坐标变换 $\boldsymbol{T}$ 或($\boldsymbol{T}$, $\boldsymbol{G}$)，将可控系统变换成简单形状的标准形．根据定理 1，因为显然反馈变换仍保存可控性，则可以考虑利用变换($\boldsymbol{T}$, $\boldsymbol{G}$, $\boldsymbol{K}$)将可控系统变换成典范形．布鲁诺夫斯基 (Brunovsky)[93]，卡尔曼 (Kalman)[94]，旺哈姆及莫尔斯(Wonham & Morse)[95]，波波夫(Popov)[96]等都各自独立地得出这种典范形.

这里仿照文献[93]，以 A-I-12 中情况 4 为基础加以说明．在情况 4 中，由 $\boldsymbol{Q}_n$ 中取出 n 个列，而被采用的仅是不能用位于其左边的列线性表出的列．这样，交换被采用列的顺序，便作出变换矩阵 $\boldsymbol{T}_{12}$(A-I-12 中(114)式)．为了以后方便起见，这里再将列的顺序加以交换，作成下列变换矩阵

$$\boldsymbol{T}_{16} \triangleq [\boldsymbol{b}_{j_1} \boldsymbol{A}\boldsymbol{b}_{j_1} \cdots \boldsymbol{A}^{\sigma_{j_1}-1}\boldsymbol{b}_{j_1} \vdots \boldsymbol{b}_{j_2} \cdots \vdots \boldsymbol{b}_{j_r} \boldsymbol{A}\boldsymbol{b}_{j_r} \cdots \boldsymbol{A}^{\sigma_{j_r}-1}\boldsymbol{b}_{j_r}] \tag{15}$$

式中(j_1, j_2, $\cdots$, j_r)是(1, 2, $\cdots$, r)的排列之一．假定(j_1, j_2, $\cdots$, j_r)的选择满足下列条件

$$\sigma_{j_1} \geqslant \sigma_{j_2} \geqslant \cdots \geqslant \sigma_{j_r} \tag{16}$$

而且，设

$$k_1 \triangleq \sigma_{j_1},\ k_2 \triangleq \sigma_{j_2},\ \cdots,\ k_r \triangleq \sigma_{j_r} \tag{17}$$

注意，因 $\boldsymbol{T}_{16}$ 和 $\boldsymbol{T}_{12}$ 只不过是顺序交换，所以只要将 σ_j 换成 k_j($j=1$, $\cdots$, r)，前章情况 4 中讲过的结果在这里都照样成立.

首先作 $\boldsymbol{T}_{16}^{-1}$，将它的第 $\sum\limits_{j=1}^{i} k_j$ 行记为 $\boldsymbol{h}_i$．而且，作

$$\boldsymbol{S} \triangleq \begin{bmatrix} \boldsymbol{h}_1 \\ \boldsymbol{h}_1\boldsymbol{A} \\ \vdots \\ \boldsymbol{h}_1\boldsymbol{A}^{k_1-1} \\ \cdots\cdots \\ \boldsymbol{h}_2 \\ \vdots \\ \cdots\cdots \\ \boldsymbol{h}_r \\ \boldsymbol{h}_r\boldsymbol{A} \\ \vdots \\ \boldsymbol{h}_r\boldsymbol{A}^{k_r-1} \end{bmatrix} \tag{18}$$

这样，变换($\boldsymbol{S}^{-1}$, $\boldsymbol{I}_r$, $\boldsymbol{0}$)将使($\boldsymbol{A}$, $\boldsymbol{B}$)变成下列典范形(A-I-12 中(63)，(124)式)

$$\boldsymbol{A}_7^c=\left[\begin{array}{c:c:c}\boldsymbol{A}_{11}'' & \cdots\cdots & \boldsymbol{A}_{1r}'' \\ \hdashline \vdots & & \vdots \\ \hdashline \boldsymbol{A}_{r1}'' & \cdots\cdots & \boldsymbol{A}_{rr}''\end{array}\right],$$

$$\boldsymbol{A}_{ii}''=\begin{bmatrix}0 & 1 & & \mathbf{0} \\ & \ddots & \ddots & \\ \mathbf{0} & & 0 & 1 \\ \cdots & \cdots & \cdots & \cdots \\ -\beta_1^{ii} & \cdots\cdots & & -\beta_{\sigma_i}^{ii}\end{bmatrix} \quad i=1,\ \cdots,\ r \tag{19}$$

$$\boldsymbol{A}_{ij}''=\begin{bmatrix} & \mathbf{0} & \\ \cdots & \cdots & \cdots \\ -\beta_1^{ij} & \cdots\cdots & -\beta_{\sigma_j}^{ij}\end{bmatrix} \quad i,\ j=1,\ \cdots,\ r;\quad i\neq j$$

$$b_{ij}=\begin{cases}0 & j>i \\ 1 & j=i \\ * & j<i\end{cases} \tag{20}$$

在这里,若作仅由 $\boldsymbol{B}_i$ 的最下行组成的 $r\times r$ 矩阵 $\boldsymbol{B}^*$

$$\boldsymbol{B}^*=\begin{bmatrix}\boldsymbol{h}_1\boldsymbol{A}^{k_1-1}\boldsymbol{B} \\ \vdots \\ \boldsymbol{h}_r\boldsymbol{A}^{k_r-1}\boldsymbol{B}\end{bmatrix}=\begin{bmatrix}b_{11} & \cdots & b_{1r} \\ \vdots & & \vdots \\ b_{r1} & \cdots & b_{rr}\end{bmatrix}, \tag{21}$$

则 $\boldsymbol{B}^*$ 是正则矩阵. (19), (20)式由变换($\boldsymbol{I}_n$, $(\boldsymbol{B}^*)^{-1}$, $\mathbf{0}$) 变成下列典范形(参看 A-I-12 中(138)式)

$$\boldsymbol{A}_{14}^c=\boldsymbol{A}_7^c \tag{22}$$

$$\boldsymbol{B}_{14}^c=\begin{bmatrix}\begin{matrix}0\\ \vdots\\ 0\\ 1\end{matrix} & & \mathbf{0} \\ & \begin{matrix}0\\ \vdots\\ 0\\ 1\end{matrix} & \\ & & \ddots \\ \mathbf{0} & & \begin{matrix}0\\ \vdots\\ 0\\ 1\end{matrix}\end{bmatrix}\begin{matrix}\}k_1\\ \\ \}k_2\\ \\ \\ \}k_r\end{matrix} \tag{23}$$

那么,这里试对(21), (22)式进行下列反馈变换

$$\boldsymbol{K}^c\triangleq\left[\begin{array}{c:c:c}\beta_1^{11}\cdots\beta_{k_1}^{11} & \cdots & \beta_1^{1r}\cdots\beta_{k_r}^{1r} \\ \cdots\cdots\cdots & & \cdots\cdots\cdots \\ \vdots & & \\ \cdots\cdots\cdots & & \cdots\cdots\cdots \\ \beta_1^{r1}\cdots\beta_{k_1}^{r1} & \cdots & \beta_1^{rr}\cdots\beta_{k_r}^{rr}\end{array}\right]. \tag{24}$$

由 $\boldsymbol{A}_{14}^c$, $\boldsymbol{B}_{14}^c$ 的形状,显然

$$
A_0 \triangleq A_{14}^c + B_{14}^c K^c = \begin{array}{cc}
\left[\begin{array}{c|c|c|c}
\begin{matrix}0&1&&\mathbf{0}\\&\ddots&\ddots&\\&&\ddots&1\\\mathbf{0}&&&0\end{matrix} & & & \mathbf{0} \\ \hline
 & \begin{matrix}0&1&&\mathbf{0}\\&\ddots&\ddots&\\&&\ddots&1\\\mathbf{0}&&&0\end{matrix} & & \\ \hline
 & & \ddots & \\ \hline
\mathbf{0} & & & \begin{matrix}0&1&&\mathbf{0}\\&\ddots&\ddots&\\&&\ddots&1\\\mathbf{0}&&&0\end{matrix}
\end{array}\right] &
\begin{array}{l} \left.\vphantom{\begin{matrix}0\\0\\0\\0\end{matrix}}\right\}k_1 \\ \left.\vphantom{\begin{matrix}0\\0\\0\\0\end{matrix}}\right\}k_2 \\ \\ \left.\vphantom{\begin{matrix}0\\0\\0\\0\end{matrix}}\right\}k_r \end{array} \\
\begin{array}{cccc} \underbrace{\qquad\qquad}_{k_1} & \underbrace{\qquad\qquad}_{k_2} & \qquad & \underbrace{\qquad\qquad}_{k_r} \end{array} &
\end{array}
$$

$$
B_0 \triangleq B_{14}^c = \left[\begin{array}{c|c|c|c}
\begin{matrix}0\\\vdots\\0\\1\end{matrix} & & & \mathbf{0} \\ \hline
 & \begin{matrix}0\\\vdots\\0\\1\end{matrix} & & \\ \hline
 & & \ddots & \\ \hline
\mathbf{0} & & & \begin{matrix}0\\\vdots\\1\end{matrix}
\end{array}\right]
\begin{array}{l} \left.\vphantom{\begin{matrix}0\\0\\0\\0\end{matrix}}\right\}k_1 \\ \left.\vphantom{\begin{matrix}0\\0\\0\\0\end{matrix}}\right\}k_2 \\ \\ \left.\vphantom{\begin{matrix}0\\0\\0\end{matrix}}\right\}k_r \end{array} \tag{25}
$$

得到仅由 0 和 1 组成的非常理想的典范形矩阵. 这是在包含反馈变换情况下所得到的最基本的典范形. 现将其推导过程再一次写出来,

$$
(A,\ B) \xrightarrow{(S^{-1},\ I_r,\ 0)} (A_{12}^c,\ B_{12}^c) \xrightarrow{(I,\ (B^*)^{-1},\ 0)} (A_{14}^c,\ B_{14}^c) \xrightarrow{(I_n,\ I_r,\ K^c)} (A_0,\ B_0) \tag{26}
$$

但是, 由 A-I-12 中(61)式, 矩阵 K^c 可用

$$
K^c = K^* S^{-1} \quad 式中\ K^* \triangleq \begin{bmatrix} h_1 A^{k_1} \\ \vdots \\ h_r A^{k_r} \end{bmatrix} \tag{27}
$$

给出. 由这个关系及(26)式可见, 对$(A,\ B)$进行下列变换

$$
(I_n,\ I_r,\ K^* S^{-1}) \circ (I_n,\ (B^*)^{-1},\ 0) \circ (S^{-1},\ I_r,\ 0) \tag{28}
$$

可得出$(A_0,\ B_0)$. 对(28)式应用(6)式, 整理后得

$$
\begin{aligned}
&(I_n,\ I_r,\ K^* S^{-1}) \circ (I_n,\ (B^*)^{-1},\ 0) \circ (S^{-1},\ I_r,\ 0) \\
&= (S^{-1},\ (B^*)^{-1},\ -(B^*)^{-1} K^*)
\end{aligned} \tag{29}
$$

即

$$
(A,\ B) \xrightarrow{(S^{-1},\ (B^*)^{-1},\ -(B^*)^{-1} K^*)} (A_0,\ B_0) \tag{30}
$$

式中 S, B^*, K^* 分别由(18), (21), (27)式给出.

这里和 A-Ⅰ-11 中一样，我们定义

$$Q_k \triangleq [B,\ AB,\ \cdots,\ A^{k-1}B] \quad k=1,\ 2,\ \cdots \tag{31}$$

$$q_k \triangleq \operatorname{rank} Q_k \tag{32}$$

若假定 $\operatorname{rank} \boldsymbol{B}=r$，$(\boldsymbol{A},\ \boldsymbol{B})$ 可控，则

$$q_1=\operatorname{rank} \boldsymbol{B}=r \tag{33}$$

$$q_\alpha=n \quad \text{式中 } \alpha \text{ 是可控性指标} \tag{34}$$

成立．但是，由变换矩阵 $\boldsymbol{T}_{16}$ 的作法可见，它包含

$$\begin{aligned}
&\boldsymbol{b}_j \text{ 形的列 } q_1=r \text{ 个，}\\
&\boldsymbol{A}\boldsymbol{b}_j \text{ 形的列 } (q_2-q_1) \text{ 个，}\\
&\qquad\vdots\\
&\boldsymbol{A}^{\alpha-1}\boldsymbol{b} \text{ 形的列 } (q_\alpha-q_{\alpha-1}) \text{ 个，}
\end{aligned} \tag{35}$$

而不包含 $\boldsymbol{A}^k\boldsymbol{b}_j(\alpha\leqslant k\leqslant n)$ 形的列．因此，对于包含在 $\boldsymbol{T}_{16}$ 中的"系列长度 k_i"，容易证明如下结果：

$$\begin{aligned}
&k_i=\alpha \quad 1\leqslant i\leqslant q_\alpha-q_{\alpha-1}\\
&k_i=\alpha-1 \quad q_\alpha-q_{\alpha-1}+1\leqslant i\leqslant q_\alpha-q_{\alpha-2}\\
&\quad\vdots\\
&k_i=1 \quad q_\alpha-q_1+1\leqslant i\leqslant q_\alpha
\end{aligned} \tag{36}$$

即若 $q_k(1\leqslant k\leqslant\alpha)$ 给定，则 $k_i(1\leqslant i\leqslant n)$ 唯一确定．而且，关于 q_k 有如下重要性质．

［引理 1］ $q_k(1\leqslant k\leqslant\alpha)$ 对于变换 $(\boldsymbol{T},\ \boldsymbol{G},\ \boldsymbol{K})$ 不变．

（证明） 和定理 1 的第二个证明方法完全相同．

由上述结果和该引理可得下列结果．

［定理］ 若 $(\boldsymbol{A},\ \boldsymbol{B})$ 可控且 $\operatorname{rank}\boldsymbol{B}=r$，则 $(\boldsymbol{A},\ \boldsymbol{B})$ 和 (25) 式形状的典范形 $(\boldsymbol{A}_0,\ \boldsymbol{B}_0)$ G 等价．典范形对于给定的 $(\boldsymbol{A},\ \boldsymbol{B})$ 唯一确定．若 $(\boldsymbol{A}_1,\ \boldsymbol{B}_1)$ 和 $(\boldsymbol{A}_2,\ \boldsymbol{B}_2)$ 是 G 等价的，则具有同一个典范形．

由该定理可见，$k_1,\ k_2,\ \cdots,\ k_r$ 是对于任意 $(\boldsymbol{T},\ \boldsymbol{G},\ \boldsymbol{K})$ 变换可以保存的重要数，称为 $(\boldsymbol{A},\ \boldsymbol{B})$ 的克罗内克指标 (Kronecker indices)．

（问题 1） 在 $\operatorname{rank}\boldsymbol{B}<r$ 的情况下，上面的标准形变成什么样？

极点配置 (Pole Assignment)

状态反馈有下面一个重要性质．

［定理 4］ 矩阵对 $(\boldsymbol{A},\ \boldsymbol{B})$ 完全可控和下列事实等价，由适当的矩阵 $\boldsymbol{K}=r\times n$，使 $\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K}$ 的特征多项式与任意配置的 n 次首一多项式一致．

$\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K}$ 的特征多项式任意选择，就是图 1 闭环系统的极点可以任意配置．不过要注意，因为我们只限于考虑 $\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{K}$ 为实矩阵的情况，特征方程式是实系数代数方程式，所以复数根必以共轭复根的形式成对地出现．也就是说，在关于复平面实轴对称的限制条件下，可以自由配置极点．

（证明） 为了简单起见，假定 $\operatorname{rank}\boldsymbol{B}=r$（因为主要特点相同）．于是由前节结果，若 $(\boldsymbol{A},\ \boldsymbol{B})$ 完全可控，则与 (25) 式*标准形 G 等价，即存在某个 $(\boldsymbol{T}_1,\ \boldsymbol{G}_1,\ \boldsymbol{K}_1)$，

$$(\boldsymbol{A},\ \boldsymbol{B}) \xrightarrow{(\boldsymbol{T}_1,\ \boldsymbol{G}_1,\ \boldsymbol{K}_1)} (\boldsymbol{A}_0,\ \boldsymbol{B}_0) \tag{37}$$

* 原文误为 (31) 式．——译者注

设给出首一多项式

$$\psi(s)=s^n+a_1s^{n-1}+\cdots+a_{n-1}s+a_n \tag{38}$$

将矩阵 $\boldsymbol{K}_0$ 选成

$$\boldsymbol{K}_0 \triangleq \left[\begin{array}{c:cccc} \boldsymbol{0} & \begin{array}{c}1\ 0\cdots 0\end{array} & & \boldsymbol{0} & \\ & \boldsymbol{0} & \ddots & \begin{array}{c}1\ 0\cdots 0\end{array} & \\ & & & & \begin{array}{c}1\ 0\cdots 0\end{array} \\ \hdashline -a_n & \cdots & \cdots & \cdots & -a_1 \end{array}\right], \quad \begin{array}{c} \underbrace{\quad}_{k_1}\ \underbrace{\quad}_{k_2}\ \cdots\ \underbrace{\quad}_{k_{r-1}}\ \underbrace{\quad}_{k_r} \end{array} \tag{39}$$

则

$$\boldsymbol{A}_0+\boldsymbol{B}_0\boldsymbol{K}_0=\begin{bmatrix} 0 & 1 & & & \boldsymbol{0} \\ & \ddots & \ddots & & \\ & \boldsymbol{0} & & \ddots & \\ & & & 0 & 1 \\ -a_n & -a_{n-1} & \cdots & -a_2 & -a_1 \end{bmatrix}, \tag{40}$$

$(\boldsymbol{A}_0+\boldsymbol{B}_0\boldsymbol{K}_0)$ 的特征多项式与 $\psi(s)$ 一致．对 $(\boldsymbol{A}_0+\boldsymbol{B}_0\boldsymbol{K}_0,\ \boldsymbol{B}_0)$ 进行 $(\boldsymbol{T}_1^{-1},\ \boldsymbol{G}_1^{-1},\ \boldsymbol{0})$ 变换，考虑到(37)式，则

$$(\boldsymbol{A}_0+\boldsymbol{B}_0\boldsymbol{K}_0,\ \boldsymbol{B}_0)\xrightarrow{(\boldsymbol{T}_1^{-1},\ \boldsymbol{G}_1^{-1},\ \boldsymbol{0})}(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K}_1+\boldsymbol{B}\boldsymbol{G}_1\boldsymbol{K}_0\boldsymbol{T}_1^{-1},\ \boldsymbol{B}) \tag{41}$$

因为(41)式的变换不改变特征多项式，则 $\boldsymbol{A}+\boldsymbol{B}(\boldsymbol{K}_1+\boldsymbol{G}_1\boldsymbol{K}_0\boldsymbol{T}_1^{-1})$ 的特征多项式仍与 $\psi(s)$ 一致．

若设 $(\boldsymbol{A},\ \boldsymbol{B})$ 不是完全可控，根据 A-I-11 中定理 5，利用适当的正则矩阵 $\boldsymbol{T}_1$ 可以变换成

$$\overline{\boldsymbol{A}}\triangleq\boldsymbol{T}_1^{-1}\boldsymbol{A}\boldsymbol{T}_1=\begin{array}{c}\begin{bmatrix}\overline{\boldsymbol{A}}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{0} & \boldsymbol{A}_{22}\end{bmatrix}\begin{array}{l}\}\ q \\ \}\ n-q,\end{array} \\ \begin{array}{cc} q & n-q \end{array}\end{array}$$

$$\overline{\boldsymbol{B}}\triangleq\boldsymbol{T}_1^{-1}\boldsymbol{B}=\begin{bmatrix}\overline{\boldsymbol{B}}_1 \\ \boldsymbol{0}\end{bmatrix}\begin{array}{l}\}\ q \\ \}\ n-q.\end{array} \tag{42}$$

设 $\boldsymbol{K}$ 为任意 $r\times n$ 矩阵，若对 $\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K}$ 进行 $\boldsymbol{T}_1$ 变换，则

$$\boldsymbol{T}_1^{-1}(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K})\boldsymbol{T}_1=\boldsymbol{T}_1^{-1}\boldsymbol{A}\boldsymbol{T}_1+\boldsymbol{T}_1^{-1}\boldsymbol{B}\boldsymbol{K}\boldsymbol{T}_1=\overline{\boldsymbol{A}}+\overline{\boldsymbol{B}}\boldsymbol{K}\boldsymbol{T}_1 \tag{43}$$

若将其中 $\boldsymbol{K}\boldsymbol{T}_1$ 分割成

$$\boldsymbol{K}\boldsymbol{T}_1=[\underbrace{\overline{\boldsymbol{K}}_1}_{q}\ \underbrace{\overline{\boldsymbol{K}}_2}_{n-q}] \tag{44}$$

则

$$\boldsymbol{T}^{-1}(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K})\boldsymbol{T}_1=\begin{bmatrix}\overline{\boldsymbol{A}}_{11}+\overline{\boldsymbol{B}}_1\overline{\boldsymbol{K}}_1 & \overline{\boldsymbol{A}}_{12}+\overline{\boldsymbol{B}}_1\overline{\boldsymbol{K}}_2 \\ \boldsymbol{0} & \overline{\boldsymbol{A}}_{22}\end{bmatrix}\triangleq\widetilde{\boldsymbol{A}} \tag{45}$$

因 $\widetilde{\boldsymbol{A}}$ 的特征多项式是 $(\overline{\boldsymbol{A}}_{11}+\overline{\boldsymbol{B}}_1\overline{\boldsymbol{K}}_1)$ 的特征多项式与 $\overline{\boldsymbol{A}}_{22}$ 的特征多项式之积，而后者与 $\boldsymbol{K}$ 无关，所以不能使 $\widetilde{\boldsymbol{A}}$ 的特征多项式与给定的任意首一多项式一致．因为 $\widetilde{\boldsymbol{A}}$ 和 $\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K}$ 的特征多项式相同，则可以说对于 $\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K}$ 也是相同的．

证明完毕

A–I–14 线性常系数系统的观测器[101a]

在前章 A–I–13 中讲过，对于线性常系数系统

$$\dot{\boldsymbol{x}}=\boldsymbol{Ax}+\boldsymbol{Bu},\ \boldsymbol{x}(t)\in R^n,\ \boldsymbol{u}(t)\in R^r,\ \boldsymbol{A}\in R^{n\times n},\ \boldsymbol{B}\in R^{n\times r}, \tag{1a}$$

$$\boldsymbol{y}=\boldsymbol{Cx},\ \boldsymbol{y}(t)\in R^m,\ \boldsymbol{C}\in R^{m\times n} \tag{1b}$$

使用状态变量反馈

$$\boldsymbol{u}=\boldsymbol{Kx}+\boldsymbol{v},\ \boldsymbol{v}(t)\in R^r,\ \boldsymbol{K}\in R^{r\times n} \tag{2}$$

时，可以使所形成的闭环系统具有各种性质．由(2)式可见，为了实现状态反馈，必须知道任意时刻的状态变量值 $\boldsymbol{x}(t)$，或者当指定反馈增益 $\boldsymbol{K}$ 时必须知道 $\boldsymbol{Kx}(t)$ 的值．

一般，所谓状态 $\boldsymbol{x}$ 是表示系统内部状态的变量(不限于物理量)，它和输入输出这样的系统外部变量不同，多半不能直接测量．

因此，采用反馈时必须考虑在不能直接获得 $\boldsymbol{x}(t)$ 值的情况下该如何办．

对于这个问题，基本上考虑如下近似方法．(i) 用被控系统的外部变量(输入和输出)产生 $\boldsymbol{x}(t)$ 的近似值(或正确值)，用它进行反馈，(ii) 考虑更直接地用输出反馈(具有和状态反馈同样效果)的补偿器．

有名的吕伯格(Luenberger)观测器[97](Observer)就是基于(i)的想法．当然，若$(\boldsymbol{A},\boldsymbol{C})$可观测，在某个有限区间内观测输出和输入，并对观测出的数值进行适当的非线性运算(参看 A–I–8 中(20)式)，便可算出起始状态 $\boldsymbol{x}(0)$，因而可算出 $\boldsymbol{x}(t)$．吕伯格考虑更简单地用线性模拟计算机处理，即使用标准形方程式表示的动力学系统近似地求 $\boldsymbol{x}(t)$．下面将要介绍这个观测器的基本内容．首先讨论观测器应该满足的充分必要条件，考察其存在性，具体地讨论几个观测器．最后，考察在通过观测器进行状态反馈的情况下闭环系统的特性如何(与可直接得到状态实现反馈的情况相比)．

观测器的基本式[98]

观测器用于产生系统(1)的状态或其线性变换 $\boldsymbol{Kx}(t)$．若令 $\boldsymbol{K}=\boldsymbol{I}_n$，则前者变成后者的一种特殊情况，所以下面我们将一般地考虑求 $\boldsymbol{Kx}(t)$ 的问题．对于观测器的构成，(i)用系统(1)的输入 $\boldsymbol{u}$ 和输出 $\boldsymbol{y}$ 作为观测器的输入，(ii)使用标准形线性常系数系统，即不使用微分器构成．在这样的条件下可以考虑

$$\dot{\boldsymbol{z}}=\boldsymbol{Fz}+\boldsymbol{Gy}+\boldsymbol{Hu} \tag{3a}$$

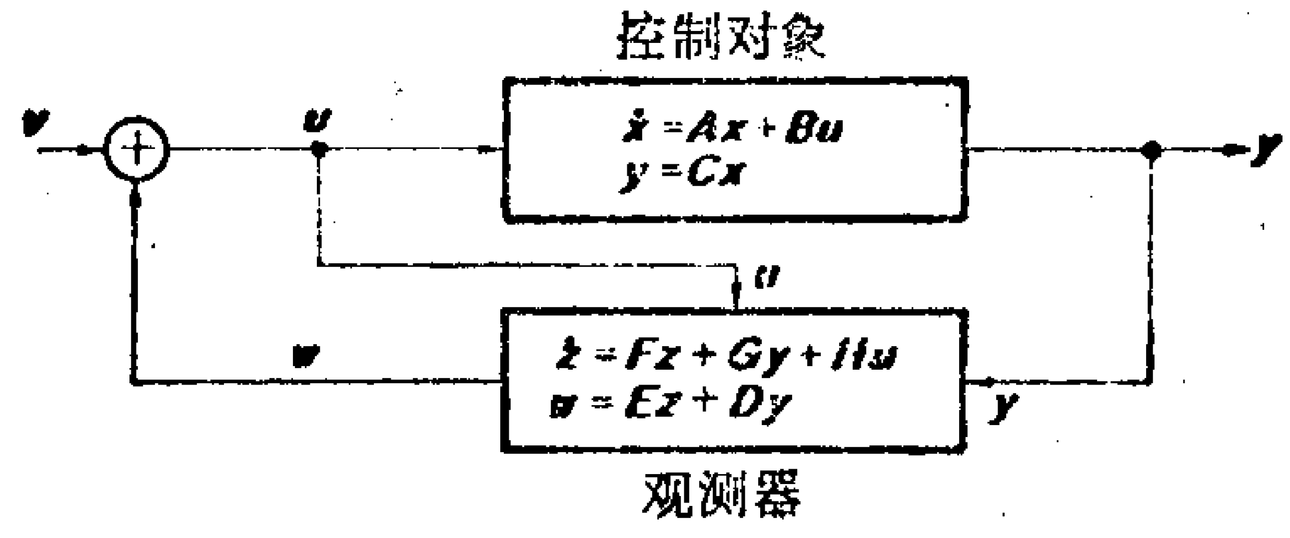


图 1　使用观测器实现状态反馈的系统$(\boldsymbol{w}(t)\to\boldsymbol{Kx}(t))$

$$\boldsymbol{w}=\boldsymbol{E}\boldsymbol{z}+\boldsymbol{D}\boldsymbol{y} \tag{3b}$$

的形式．式中 $\boldsymbol{z}$ 是观测器的状态，其维数是 $\boldsymbol{z}(t)\in R^p$，在该阶段假定 p 的值未定．$\boldsymbol{w}(t)$ 是输出，$\boldsymbol{w}(t)\in R^r$ 给出 $\boldsymbol{K}\boldsymbol{x}(t)$ 的近似值．

［定义 1］ 在系统(1)，(3)中，不论 $\boldsymbol{x}(0)$，$\boldsymbol{z}(0)$ 的值及 $\boldsymbol{u}(\cdot)$ 如何，当 $\lim\limits_{t\to\infty}[\boldsymbol{w}(t)-\boldsymbol{K}\boldsymbol{x}(t)]=\boldsymbol{0}$ 成立时，称系统(3)为(1)的观测器(或 $\boldsymbol{K}\boldsymbol{x}$ 观测器)．特别是，在 $\boldsymbol{K}=\boldsymbol{I}_n$ 时称为状态观测器．

那么可以预料，为使 $\boldsymbol{w}(t)$ 收敛于 $\boldsymbol{K}\boldsymbol{x}(t)$，观测器的状态 $\boldsymbol{z}(t)$ 必须收敛于 $\boldsymbol{x}(t)$ 的某个线性变换 $\boldsymbol{T}\boldsymbol{x}(t)$．事实上，以后可以证明这个结论是正确的，现在我们先来求其成立的条件．

［引理 1］[98] 设 $(\boldsymbol{A},\boldsymbol{B})$ 为可控对．在系统(1)，(3)中，与 $\boldsymbol{x}(0)$ 及 $\boldsymbol{z}(0)$ 无关，$\lim\limits_{t\to\infty}(\boldsymbol{z}(t)-\boldsymbol{T}\boldsymbol{x}(t))=\boldsymbol{0}$ (对于某个给定的 $\boldsymbol{T}\in R^{p\times n}$)的充分必要条件是

(i) $$\mathrm{Re}\lambda_i(\boldsymbol{F})<0,\ i=1,\ 2,\ \cdots,\ p \tag{4a}$$

(ii) $$\boldsymbol{T}\boldsymbol{A}-\boldsymbol{F}\boldsymbol{T}=\boldsymbol{G}\boldsymbol{C} \tag{4b}$$

(iii) $$\boldsymbol{H}=\boldsymbol{T}\boldsymbol{B} \tag{4c}$$

成立．

(证明概略) 充分性：

设 $\tilde{\boldsymbol{z}}\triangleq\boldsymbol{z}-\boldsymbol{T}\boldsymbol{x}$，由(1)，(3)式得

$$\dot{\tilde{\boldsymbol{z}}}=\boldsymbol{F}\tilde{\boldsymbol{z}}-(\boldsymbol{T}\boldsymbol{A}-\boldsymbol{F}\boldsymbol{T}-\boldsymbol{G}\boldsymbol{C})\boldsymbol{x}+(\boldsymbol{H}-\boldsymbol{T}\boldsymbol{B})\boldsymbol{u}. \tag{5}$$

若(4b)，(4c)式成立，则变成关于 $\tilde{\boldsymbol{z}}$ 的自由运动方程式 $\dot{\tilde{\boldsymbol{z}}}=\boldsymbol{F}\tilde{\boldsymbol{z}}$．而且由(4a)式得 $\tilde{\boldsymbol{z}}\to\boldsymbol{0}$．

必要性：对于任意 $\boldsymbol{x}(0)$，$\boldsymbol{z}(0)$，$\boldsymbol{u}(\cdot)$，设 $\tilde{\boldsymbol{z}}(t)\to\boldsymbol{0}$．特别是，因为若 $\boldsymbol{x}(0)=\boldsymbol{0}$，$\boldsymbol{u}(t)=\boldsymbol{0}$，$t\geqslant0$，则 $\boldsymbol{x}(t)=\boldsymbol{0}$，$t\geqslant0$，所以 $\tilde{\boldsymbol{z}}(t)=\boldsymbol{z}(t)$，对于 $\tilde{\boldsymbol{z}}$，$\dot{\tilde{\boldsymbol{z}}}=\boldsymbol{F}\tilde{\boldsymbol{z}}$ 成立．因为对于任意 $\tilde{\boldsymbol{z}}(0)\ (=\boldsymbol{z}(0))$，$\tilde{\boldsymbol{z}}(t)\to\boldsymbol{0}$，则(4a)式必须成立．

此外，若(4c)式不成立，则(5)式第 3 项可设定得任意大．因此，为了使 $\tilde{\boldsymbol{z}}(t)\to\boldsymbol{0}$，(4c)必须成立．最后，若(4b)不成立，因系统(1)是可控的，则由适当的输入来控制 $\boldsymbol{x}(t)$，可使(5)式第 2 项任意大．为了使 $\tilde{\boldsymbol{z}}(t)\to\boldsymbol{0}$，因此(4b)式仍然必须成立．

证明完毕

下面我们来证明，为了使(3)式输出 $\boldsymbol{w}(t)$ 收敛于 $\boldsymbol{K}\boldsymbol{x}(t)$，实际上 $\boldsymbol{z}(t)\to\boldsymbol{T}\boldsymbol{x}(t)$ 是必要的．

［引理 2］[98] 设 $(\boldsymbol{A},\boldsymbol{B})$ 可控，$(\boldsymbol{E},\boldsymbol{F})$ 可观测[1)]．系统(3)为(1)的观测器的充分必要条件是，对于某个 $\boldsymbol{T}$，$\boldsymbol{z}(t)\to\boldsymbol{T}\boldsymbol{x}(t)$ (即(4a)，(4b)，(4c)式成立)，而且

$$\boldsymbol{K}=\boldsymbol{E}\boldsymbol{T}+\boldsymbol{D}\boldsymbol{C} \tag{6}$$

成立．

(证明概略) 充分性：若 $\boldsymbol{z}(t)\to\boldsymbol{T}\boldsymbol{x}(t)$，且(6)式成立，由(3b)式，则

$$\boldsymbol{w}(t)=\boldsymbol{E}\boldsymbol{z}(t)+\boldsymbol{D}\boldsymbol{y}(t)\longrightarrow(\boldsymbol{E}\boldsymbol{T}+\boldsymbol{D}\boldsymbol{C})\boldsymbol{x}(t)=\boldsymbol{K}\boldsymbol{x}(t).$$

必要性：假定系统(3)是(1)的观测器．设输入 $\boldsymbol{u}(\cdot)$ 选自解析函数的集合[2)]，这意味

1) 假定 $(\boldsymbol{E},\boldsymbol{F})$ 可观测并不失去普遍性．若可观测性不成立，则可换成 A-I-11 中讲过的等价系统．

2) 一般对于解析函数 f，既使 $f(t)\to0$，$t\to\infty$，也不一定 $\dot{f}(t)\to0$，$t\to\infty$．但是在常系数标准形状态方程的零输入响应(由起始状态决定的响应)中，上述性质成立．因为这里假定系统(1)可控，所以在以下证明中可以只考虑将 $\boldsymbol{x}$ 反馈得到的一类解析输入 $\boldsymbol{u}(t)=\boldsymbol{L}\boldsymbol{x}(t)$．这是由池田氏(神户大学)所指出．

着(1), (3)式的解 $\boldsymbol{w}$, $\boldsymbol{z}$, $\boldsymbol{x}$ 均为解析函数, $\boldsymbol{w}(t)\to\boldsymbol{Kx}(t)$, 以及其微分 $\boldsymbol{w}^{(j)}(t)\to\boldsymbol{Kx}^{(j)}(t)$, $j=1, 2, \cdots$均成立. 实际上计算一下微分, 对于 $j=1$, 得

$$\boldsymbol{w}^{(1)}(t)-\boldsymbol{Kx}^{(1)}(t)=\boldsymbol{EFz}(t)+\boldsymbol{M}_1\boldsymbol{x}(t)+\boldsymbol{N}_1\boldsymbol{u}_1(t).$$

$\boldsymbol{M}$, $\boldsymbol{N}$ 由系统(1), (3)的系数矩阵和 $\boldsymbol{K}$ 决定, 为了简单起见, 用 $\boldsymbol{M}_1$, $\boldsymbol{N}_1$ 表示.

在这里与 $\boldsymbol{u}(\cdot)$ 无关, 为使 $\boldsymbol{w}^{(1)}(t)\to\boldsymbol{Kx}^{(1)}(t)$, 必须 $\boldsymbol{N}_1=\boldsymbol{0}$. 以下对于 $j=2, 3, \cdots$也是一样. 将对于 $j=0, 1, 2, \cdots, p-1$ 的关系式整理后可以写成

$$\lim_{t\to\infty}\left\{\begin{bmatrix}\boldsymbol{w}(t)\\ \boldsymbol{w}^{(1)}(t)\\ \vdots\\ \boldsymbol{w}^{(p-1)}(t)\end{bmatrix}-\begin{bmatrix}\boldsymbol{Kx}(t)\\ \boldsymbol{Kx}^{(1)}(t)\\ \vdots\\ \boldsymbol{Kx}^{(p-1)}(t)\end{bmatrix}\right\}=\lim_{t\to\infty}\{\boldsymbol{Rz}(t)-\boldsymbol{Mx}(t)\}=\boldsymbol{0}, \tag{7}$$

式中

$$\boldsymbol{R}\triangleq\begin{bmatrix}\boldsymbol{E}\\ \boldsymbol{EF}\\ \vdots\\ \boldsymbol{EF}^{p-1}\end{bmatrix}, \quad \boldsymbol{M}\triangleq\begin{bmatrix}\boldsymbol{M}_0\\ \boldsymbol{M}_1\\ \vdots\\ \boldsymbol{M}^{p-1}\end{bmatrix}.$$

$\boldsymbol{M}_i(i=0, 1, \cdots, p-1)$ 是由系统(1), (3)的系数矩阵 $\boldsymbol{A}$, $\boldsymbol{C}$, $\boldsymbol{D}$, $\boldsymbol{E}$, $\boldsymbol{G}$ 及 $\boldsymbol{K}$ 决定的矩阵. 因$(\boldsymbol{E}, \boldsymbol{F})$可观测, 则 $\boldsymbol{R}$ 具有最大秩 p. 因此, $p\times p$ 矩阵 $\boldsymbol{R}^T\boldsymbol{R}$ 是正则的. 令 $\boldsymbol{T}\triangleq(\boldsymbol{R}^T\boldsymbol{R})^{-1}\boldsymbol{R}^T\boldsymbol{M}$, 则(7)式意味着 $\boldsymbol{z}(t)\to\boldsymbol{Tx}(t)$.

那么, 若假定 $\boldsymbol{z}(t)\to\boldsymbol{Tx}(t)$, 利用(3b)式则得

$$\lim_{t\to\infty}\boldsymbol{E}(\boldsymbol{z}(t)-\boldsymbol{Tx}(t))=\lim_{t\to\infty}\{\boldsymbol{w}(t)-(\boldsymbol{DC}+\boldsymbol{ET})\boldsymbol{x}(t)\}=\boldsymbol{0}.$$

因 $\boldsymbol{w}(t)\to\boldsymbol{Kx}(t)$, 则这意味着

$$\lim_{t\to\infty}(\boldsymbol{K}-\boldsymbol{DC}-\boldsymbol{ET})\boldsymbol{x}(t)=\boldsymbol{0}. \tag{8}$$

若 $\boldsymbol{K}=\boldsymbol{DC}+\boldsymbol{ET}$ 不成立, 用适当的(解析的)输入来控制 $\boldsymbol{x}(t)$, 可以使(8)式不成立. 因此, (6)式必须成立.

证明完毕

于是由引理 1, 2 可见, 对于可控系统(1), 为使系统(3)是定义 1 意义上的观测器, 其充分必要条件是, 对于某个 $\boldsymbol{T}$, (4), (6)式必须成立. 当给定系统 (1), $\boldsymbol{K}$ 指定时, 为了构成观测器, 找出满足这些式子的矩阵 $\boldsymbol{F}$, $\boldsymbol{G}$, $\boldsymbol{E}$, $\boldsymbol{D}$ 及 $\boldsymbol{T}$ 即可.

$\boldsymbol{T}$ 求出后, 设 $\boldsymbol{H}=\boldsymbol{TB}$, 则 $\boldsymbol{H}$ 唯一确定, 而且(6)式可以写成下列分块矩阵的形式

$$\boldsymbol{K}=[\boldsymbol{E}, \boldsymbol{D}]\begin{bmatrix}\boldsymbol{T}\\ \boldsymbol{C}\end{bmatrix}. \tag{9}$$

满足(9)式的 $\boldsymbol{E}$, $\boldsymbol{D}$ 存在等价于

$$\text{“}\boldsymbol{K}\text{ 的各行}=\boldsymbol{T}, \boldsymbol{C}\text{ 各行的线性组合”} \tag{10}$$

因此, 观测器存在的充分必要条件是, 对于某个 $\boldsymbol{F}$, $\boldsymbol{G}$, $\boldsymbol{T}$, (4)及(10)成立. 若$(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$及 $\boldsymbol{K}$ 满足各种条件, 当这样的 $\boldsymbol{F}$, $\boldsymbol{G}$, $\boldsymbol{T}$ 存在, 而且不是唯一存在时, 观测器的维数能小到什么程度. 关于这个问题没有完全解决, 下面谈一下这方面的主要结果.

状态观测器

吕伯格也曾证明过, 若$(\boldsymbol{A}, \boldsymbol{C})$可观测, 则始终有状态观测器$(\boldsymbol{K}=\boldsymbol{I}_n)$存在. 现在假

定 $p=n$, 而且若令 $\boldsymbol{T}=\boldsymbol{I}_n$, 则 $\boldsymbol{D}=\boldsymbol{0}$, $\boldsymbol{E}=\boldsymbol{I}_n$, $\boldsymbol{H}=\boldsymbol{B}$, $\boldsymbol{F}=\boldsymbol{A}-\boldsymbol{GC}$ 中的 $\boldsymbol{D}$, $\boldsymbol{E}$, $\boldsymbol{H}$, $\boldsymbol{F}$ 满足(4b), (4c), (6)式. 设$(\boldsymbol{A}, \boldsymbol{C})$可观测. 因 $(\boldsymbol{A}^T, \boldsymbol{C}^T)$ 可控, 如前章所述, 适当地设定 $\boldsymbol{G}$, 则 $\boldsymbol{F}^T=\boldsymbol{A}^T-\boldsymbol{C}^T\boldsymbol{G}^T$ 的特征值(即 $\boldsymbol{F}$ 的特征值)可以自由配置. 当然应使 $\mathrm{Re}\lambda_i(\boldsymbol{F})<0$, $i=1, 2, \cdots, n$, 即 $\boldsymbol{G}$ 的选择应满足(4a)式. 将以上整理得如下定理:

[定理 1][99] 若$(\boldsymbol{A}, \boldsymbol{C})$可观测,则对于系统(1)可以构成

$$\dot{\boldsymbol{z}}=(\boldsymbol{A}-\boldsymbol{GC})\boldsymbol{z}+\boldsymbol{Gy}+\boldsymbol{Bu} \tag{11a}$$

$$\boldsymbol{w}=\boldsymbol{z} \tag{11b}$$

的 n 维状态观测器. 而且, $(\boldsymbol{A}-\boldsymbol{GC})$的特征值在左半平面内可以任意配置(当然是以共轭复根的形式). 若$(\boldsymbol{A}, \boldsymbol{C})$不可观测,则不能任意配置.

(注意)

"因 $\boldsymbol{T}=\boldsymbol{I}_n$, $\boldsymbol{w}=\boldsymbol{z}$, 根据(5)式, 则观测误差 $\tilde{\boldsymbol{z}}=\boldsymbol{z}-\boldsymbol{x}$ 的动态变化过程可由 $\dot{\tilde{\boldsymbol{z}}}=(\boldsymbol{A}-\boldsymbol{GC})\tilde{\boldsymbol{z}}$ 规定. 如上所述, 设定适当的 $\boldsymbol{G}$, 可以给予对应于 $\tilde{\boldsymbol{z}}(0)$ 的误差 $\tilde{\boldsymbol{z}}(t)$ 以任意的衰减率."

那么, 在定理 1 中$(\boldsymbol{A}, \boldsymbol{C})$可观测是状态观测器存在的充分条件. 作为其中的一个例子, 介绍了 n 维状态观测器. 实际上可以证明[99−101], 要构成状态观测器不一定要 n 维, 维数可以更低,事实上可用 $p=(n-\mathrm{rank}\,\boldsymbol{C})$维构成,而且这是理论上最小值.

不失去普遍性,现在我们假定 $\boldsymbol{C}\in R^{m\times n}$ 有最大秩 m, 设变换矩阵 $\boldsymbol{P}\in R^{n\times n}$, 考虑

$$\boldsymbol{P}\triangleq\begin{bmatrix}\boldsymbol{C}\\ \boldsymbol{C}_2\end{bmatrix}\begin{matrix}\}\ m\\ \}\ n-m,\end{matrix}$$

由 $\boldsymbol{P}^{-1}$ 将$(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$变换成代数等价的$(\overline{\boldsymbol{A}}, \overline{\boldsymbol{B}}, \overline{\boldsymbol{C}})$. 其中 $\boldsymbol{C}_2$ 是在 $\boldsymbol{P}$ 为正则矩阵的范围内选择的任意矩阵. 在变换后得到的系统中

$$\bar{\boldsymbol{x}}\triangleq\boldsymbol{Px}=\begin{bmatrix}\boldsymbol{Cx}\\ \boldsymbol{C}_2\boldsymbol{x}\end{bmatrix}\triangleq\begin{bmatrix}\boldsymbol{y}\\ \bar{\boldsymbol{x}}_2\end{bmatrix}\begin{matrix}\}\ m\\ \}\ n-m\end{matrix} \tag{12a}$$

$$\overline{\boldsymbol{A}}\triangleq\boldsymbol{PAP}^{-1}\triangleq\begin{bmatrix}\overline{\boldsymbol{A}}_{11} & \overline{\boldsymbol{A}}_{12}\\ \underbrace{\overline{\boldsymbol{A}}_{21}}_{m} & \underbrace{\overline{\boldsymbol{A}}_{22}}_{n-m}\end{bmatrix}\begin{matrix}\}\ m\\ \}\ n-n\end{matrix} \tag{12b}$$

$$\overline{\boldsymbol{B}}\triangleq\boldsymbol{PB}\triangleq\begin{bmatrix}\overline{\boldsymbol{B}}_1\\ \overline{\boldsymbol{B}}_2\end{bmatrix}\begin{matrix}\}\ m\\ \}\ n-m\end{matrix} \tag{12c}$$

$$\overline{\boldsymbol{C}}\triangleq\boldsymbol{CP}^{-1}=[\ \underbrace{\boldsymbol{I}_m}_{m}\ \ \underbrace{\boldsymbol{0}}_{n-m}\ ]. \tag{12d}$$

对于(12)式,首先下列结果成立.

[引理 3][100]

若$(\boldsymbol{A}, \boldsymbol{C})$可观测,则$(\overline{\boldsymbol{A}}_{22}, \overline{\boldsymbol{A}}_{12})$亦可观测.

(证明)

因$(\overline{\boldsymbol{A}}, \overline{\boldsymbol{C}})$可观测,则其可观测性矩阵 $\overline{\boldsymbol{R}}$ 的秩为 n, 即

$$\mathrm{rank}\,\overline{\boldsymbol{R}}\triangleq\mathrm{rank}\begin{bmatrix}\boldsymbol{I}_m & \boldsymbol{0}\\ \overline{\boldsymbol{A}}_{11} & \overline{\boldsymbol{A}}_{12}\\ \overline{\boldsymbol{A}}_{11}^2+\overline{\boldsymbol{A}}_{12}\overline{\boldsymbol{A}}_{21} & \overline{\boldsymbol{A}}_{11}\overline{\boldsymbol{A}}_{12}+\overline{\boldsymbol{A}}_{12}\overline{\boldsymbol{A}}_{22}\\ \cdots\cdots & \cdots\cdots\end{bmatrix}=n$$

因对 $\tilde{\boldsymbol{R}}$ 进行基本行变换其秩不变，则容易证明

$$\operatorname{rank} \overline{\boldsymbol{R}}=\operatorname{rank}\begin{bmatrix} \boldsymbol{I}_m & \boldsymbol{0} \\ \boldsymbol{0} & \overline{\boldsymbol{A}}_{12} \\ \boldsymbol{0} & \overline{\boldsymbol{A}}_{12}\overline{\boldsymbol{A}}_{22} \\ \vdots & \vdots \\ \boldsymbol{0} & \overline{\boldsymbol{A}}_{12}\overline{\boldsymbol{A}}_{22}^{n-1} \end{bmatrix} \tag{13}$$

由(13)式可见

$$\operatorname{rank}\begin{bmatrix} \overline{\boldsymbol{A}}_{12} \\ \overline{\boldsymbol{A}}_{12}\overline{\boldsymbol{A}}_{22} \\ \vdots \\ \overline{\boldsymbol{A}}_{12}\overline{\boldsymbol{A}}_{22}^{n-1} \end{bmatrix}=n-m$$

即 $(\boldsymbol{A}_{22}, \overline{\boldsymbol{A}}_{12})$ 可观测.

证明完毕

那末，我们假定 $(\boldsymbol{A}, \boldsymbol{C})$ 可观测，考虑对于系统 $(\overline{\boldsymbol{A}}, \overline{\boldsymbol{B}}, \overline{\boldsymbol{C}})$ 构成 $p=n-m$ 维状态观测器. 在这种情况下，因为状态 $\overline{\boldsymbol{x}}$ 的 m 个分量作为输出 $\boldsymbol{y}$ (参看(12a)式)可以直接给出，所以只要按照(11)式作出对应于剩下的 $(n-m)$ 个分量 $\overline{\boldsymbol{x}}_2$ 的状态观测器即可. 对应于 $\overline{\boldsymbol{x}}_2$ 的状态方程式为

$$\dot{\overline{\boldsymbol{x}}}_2=\overline{\boldsymbol{A}}_{22}\overline{\boldsymbol{x}}_2+\overline{\boldsymbol{A}}_{21}\boldsymbol{y}+\overline{\boldsymbol{B}}_2\boldsymbol{u}, \tag{14a}$$

但没有相当的输出方程式. 因此，将关于 $\overline{\boldsymbol{x}}$ 的最初 m 个分量 $\boldsymbol{y}$ 的状态方程式 $\dot{\boldsymbol{y}}=\boldsymbol{A}_{11}\boldsymbol{y}+\boldsymbol{A}_{12}\overline{\boldsymbol{x}}_2+\overline{\boldsymbol{B}}_1\boldsymbol{u}$ 改写成

$$\dot{\boldsymbol{y}}-\overline{\boldsymbol{A}}_{11}\boldsymbol{y}-\overline{\boldsymbol{B}}_1\boldsymbol{u}=\boldsymbol{A}_{12}\overline{\boldsymbol{x}}_2 \tag{14b}$$

后，看作输出方程式. 参照(11)式，则对应于系统(14)的状态观测器可以写成

$$\dot{\hat{\boldsymbol{z}}}=(\overline{\boldsymbol{A}}_{22}-\overline{\boldsymbol{G}}_2\overline{\boldsymbol{A}}_{12})\hat{\boldsymbol{z}}+\overline{\boldsymbol{G}}_2(\dot{\boldsymbol{y}}-\overline{\boldsymbol{A}}_{11}\boldsymbol{y}-\overline{\boldsymbol{B}}_1\boldsymbol{u})+\overline{\boldsymbol{A}}_{21}\boldsymbol{y}+\overline{\boldsymbol{B}}_2\boldsymbol{u} \tag{15a}$$

$$\overline{\boldsymbol{w}}_2=\hat{\boldsymbol{z}} \tag{15b}$$

象这样，形成 $\dot{\boldsymbol{y}}$ 就需要微分器，故仍使用在 A-I-2 中已经熟悉的方法. 即，设 $\boldsymbol{z}\triangleq\hat{\boldsymbol{z}}-\overline{\boldsymbol{G}}_2\boldsymbol{y}$，则可以等价变换成下列标准形

$$\begin{aligned}\dot{\boldsymbol{z}}=&(\overline{\boldsymbol{A}}_{22}-\overline{\boldsymbol{G}}_2\overline{\boldsymbol{A}}_{12})\boldsymbol{z}+\{(\overline{\boldsymbol{A}}_{22}-\overline{\boldsymbol{G}}_2\overline{\boldsymbol{A}}_{12})\overline{\boldsymbol{G}}_2+\overline{\boldsymbol{A}}_{21}-\overline{\boldsymbol{G}}_2\overline{\boldsymbol{A}}_{11}\}\boldsymbol{y}\\&+(\overline{\boldsymbol{B}}_2-\overline{\boldsymbol{G}}_2\overline{\boldsymbol{B}}_1)\boldsymbol{u}\end{aligned} \tag{16a}$$

$$\overline{\boldsymbol{w}}_2=\boldsymbol{z}+\overline{\boldsymbol{G}}_2\boldsymbol{y} \tag{16b}$$

(16)式给出对应于 $\overline{\boldsymbol{x}}_2$ 的 $(n-m)$ 维状态观测器. 注意，特别是因 $(\overline{\boldsymbol{A}}_{22}, \overline{\boldsymbol{A}}_{12})$ 可观测(引理3)，则 $(\overline{\boldsymbol{A}}_{22}-\overline{\boldsymbol{G}}_2\overline{\boldsymbol{A}}_{12})$ 的特征值可以任意配置.

对应于系统 $(\overline{\boldsymbol{A}}, \overline{\boldsymbol{B}}, \overline{\boldsymbol{C}})$ 的状态观测器，我们用(16a)式和输出方程式

$$\overline{\boldsymbol{w}}=\begin{bmatrix}\boldsymbol{y}\\ \overline{\boldsymbol{w}}_2\end{bmatrix}=\begin{bmatrix}\boldsymbol{0}\\ \boldsymbol{I}_{n-m}\end{bmatrix}\boldsymbol{z}+\begin{bmatrix}\boldsymbol{I}_m\\ \overline{\boldsymbol{G}}_2\end{bmatrix}\boldsymbol{y} \tag{17}$$

表示. 将这个 $(n-m)$ 维状态观测器(16a)，(17)式改写成 $\dot{\boldsymbol{z}}=\boldsymbol{F}\boldsymbol{z}+\boldsymbol{G}\boldsymbol{y}+\boldsymbol{H}\boldsymbol{u}$, $\overline{\boldsymbol{w}}=\overline{\boldsymbol{E}}\boldsymbol{z}+\overline{\boldsymbol{D}}\boldsymbol{y}$ 时可以看到，这些系数矩阵 $\boldsymbol{F}, \boldsymbol{G}, \boldsymbol{H}, \overline{\boldsymbol{E}}, \overline{\boldsymbol{D}}$ 对于下列

$$\overline{\boldsymbol{T}}=[-\overline{\boldsymbol{G}}_2\ \boldsymbol{I}_{n-m}]$$

的矩阵 $\overline{\boldsymbol{T}}\in R^{(n-m)\times n}$，满足基本式(4)，(6)，即 $\overline{\boldsymbol{T}}\,\overline{\boldsymbol{A}}-\boldsymbol{F}\overline{\boldsymbol{T}}=\boldsymbol{G}\overline{\boldsymbol{C}}$, $\boldsymbol{H}=\overline{\boldsymbol{T}}\overline{\boldsymbol{B}}$, $\boldsymbol{I}_n=\overline{\boldsymbol{E}}\,\overline{\boldsymbol{T}}+\overline{\boldsymbol{D}}\overline{\boldsymbol{C}}$ 成立. 对于变换以前的系统 $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ 因 $(\overline{\boldsymbol{A}}, \overline{\boldsymbol{B}}, \overline{\boldsymbol{C}})$，的状态 $\overline{\boldsymbol{x}}$ 和该系统的状态 $\boldsymbol{x}$

之间存在 $\boldsymbol{x}=\boldsymbol{P}^{-1}\bar{\boldsymbol{x}}$ 的关系，则仅将(17)式的输出方程式变换成 $\boldsymbol{w}=\boldsymbol{P}^{-1}\bar{\boldsymbol{w}}$ 后即可得到其状态观测器．即，系统 $(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{C})$ 的状态观测器可以用(16a)式和输出方程式

$$\boldsymbol{w}=\boldsymbol{E}\boldsymbol{z}+\boldsymbol{D}\boldsymbol{y},\ \boldsymbol{E}\triangleq\boldsymbol{P}^{-1}\bar{\boldsymbol{E}},\ \boldsymbol{D}\triangleq\boldsymbol{P}^{-1}\bar{\boldsymbol{D}} \tag{18}$$

表示．(16a)，(18)两式的系数矩阵 $\boldsymbol{F},\ \boldsymbol{G},\ \boldsymbol{H},\ \boldsymbol{E},\ \boldsymbol{D}$，对于 $\boldsymbol{T}\triangleq\bar{\boldsymbol{T}}\boldsymbol{P}$ 满足基本式，即 $\boldsymbol{T}\boldsymbol{A}-\boldsymbol{F}\boldsymbol{T}=\boldsymbol{G}\boldsymbol{C},\ \boldsymbol{H}=\boldsymbol{T}\boldsymbol{B},\ \boldsymbol{I}_n=\boldsymbol{E}\boldsymbol{T}+\boldsymbol{D}\boldsymbol{C}$ 成立．

那么，实际上该 $(n-m)$ 维状态观测器(16a)，(18)式具有最小维数．其原因是，设 $\boldsymbol{K}=\boldsymbol{I}_n$，因(6)式成立，则必须

$$\operatorname{rank}\begin{bmatrix}\boldsymbol{T}\\ \boldsymbol{C}\end{bmatrix}=n.$$

在这种情况下，因 $\operatorname{rank}\boldsymbol{T}\geqslant n-\operatorname{rank}\boldsymbol{C}\geqslant n-m$ 必须成立，则状态观测器维数的最小值是 $n-m$．将以上整理得如下定理．

［定理 2］ 若 $(\boldsymbol{A},\ \boldsymbol{C})$ 可观测，$\operatorname{rank}\boldsymbol{C}=m$，则对系统(1)可以构成 $(n-m)$ 维状态观测器 $\dot{\boldsymbol{z}}=\boldsymbol{F}\boldsymbol{z}+\boldsymbol{G}\boldsymbol{y}+\boldsymbol{H}\boldsymbol{u},\ \boldsymbol{w}=\boldsymbol{E}\boldsymbol{z}+\boldsymbol{D}\boldsymbol{y}$．而且，$\boldsymbol{F}$ 的特征值在左半平面内可任意配置．若再假定 $(\boldsymbol{A},\ \boldsymbol{B})$ 可控，则该状态观测器有最小维数．

（注意）

和定理 1 的情况一样，使用对应于系统 $(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{C})$ 的 $(n-m)$ 维状态观测器(16a)，(18)式时的误差 $\boldsymbol{e}=\boldsymbol{x}-\boldsymbol{w}$，也可以给出任意的衰减率．即

$$\boldsymbol{e}=\boldsymbol{x}-\boldsymbol{w}=\boldsymbol{P}^{-1}(\bar{\boldsymbol{x}}-\bar{\boldsymbol{w}}_2)=\boldsymbol{P}^{-1}\begin{bmatrix}\boldsymbol{0}\\ \bar{\boldsymbol{x}}_2-\bar{\boldsymbol{w}}_2\end{bmatrix}.$$

而且，由(14)，(15)式显而易见，可以表示成

$$\dot{\bar{\boldsymbol{x}}}_2-\dot{\bar{\boldsymbol{w}}}_2=(\bar{\boldsymbol{A}}_{22}-\bar{\boldsymbol{G}}_2\bar{\boldsymbol{A}}_{12})(\bar{\boldsymbol{x}}_2-\bar{\boldsymbol{w}}_2).$$

因状态方程式左乘以 $\boldsymbol{K}$ 显然变成 $\boldsymbol{K}\boldsymbol{x}$-观测器，根据定理 2，则对于任意 $\boldsymbol{K}$ 都可以构成 $(n-m)$ 维 $\boldsymbol{K}\boldsymbol{x}$-观测器．但是，一般 $(n-m)$ 不一定是 $\boldsymbol{K}\boldsymbol{x}$-观测器的最小维数．特别是，在 $\boldsymbol{K}$ 是行向量，即 $\boldsymbol{K}\boldsymbol{x}$ 是纯量的情况下(这时 $\boldsymbol{K}\boldsymbol{x}$-观测器称为泛函观测器)，对于任意的 $\boldsymbol{K}$，当令 ν 为 $(\boldsymbol{A},\ \boldsymbol{C})$ 的可观测性指标时，都可以构成 $(\nu-1)$ 维观测器．注意，当 $(\boldsymbol{A},\ \boldsymbol{C})$ 可观测时，一般 $\nu-1\leqslant n-m$ 成立．

闭环系统的特征值

最后我们来讨论一下，当用观测器构成闭环系统时全体特征值是怎样分布的．若反馈量 $\boldsymbol{K}\boldsymbol{x}$ 可以直接得到，则 A-I-13 中讲过的闭环系统的状态方程式变成 $\dot{\boldsymbol{x}}=(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K})\times\boldsymbol{x}+\boldsymbol{B}\boldsymbol{v}$．即闭环系统的特征值是 $\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K}$ 的特征值．但是，$\boldsymbol{K}\boldsymbol{x}$ 不能直接利用，我们将以(3)式观测器作为要使用的观测器．实际上在这种情况下，关于闭环系统的特征值下述“分离性”定理成立．这在设计上是非常方便的．

［定理 3］[198] 对于系统(1)，在用满足(4b)，(4c)，(6)式的观测器(3)构成闭环系统的情况下，闭环系统的特征值由不用观测器而直接将 $\boldsymbol{K}\boldsymbol{x}$ 反馈时的特征值(即 $\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K}$ 的特征值)和观测器的特征值(即 $\boldsymbol{F}$ 的特征值)组成．

（证明） 将(1)，(3)式合起来，闭环系统的状态方程式可以用下式表示

$$\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{u},\ \ \boldsymbol{y}=\boldsymbol{C}\boldsymbol{x}$$
$$\dot{\boldsymbol{z}}=\boldsymbol{F}\boldsymbol{z}+\boldsymbol{G}\boldsymbol{y}+\boldsymbol{H}\boldsymbol{u}$$

$$\boldsymbol{u}=\boldsymbol{E}\boldsymbol{z}+\boldsymbol{D}\boldsymbol{y}+\boldsymbol{v}.$$

这里若再对闭环系统的状态 $\boldsymbol{x}$, $\boldsymbol{z}$ 进行下列正则变换

$$\begin{bmatrix} \boldsymbol{x} \\ \tilde{\boldsymbol{z}} \end{bmatrix}=\begin{bmatrix} \boldsymbol{I}_n & \boldsymbol{0} \\ -\boldsymbol{T} & \boldsymbol{I}_p \end{bmatrix}\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{z} \end{bmatrix},$$

对于新的状态,因(4b), (4c), (6)式成立,则

$$\begin{bmatrix} \dot{\boldsymbol{x}} \\ \dot{\tilde{\boldsymbol{z}}} \end{bmatrix}=\begin{bmatrix} \boldsymbol{A}+\boldsymbol{B}\boldsymbol{K} & \boldsymbol{B}\boldsymbol{E} \\ \boldsymbol{0} & \boldsymbol{F} \end{bmatrix}\begin{bmatrix} \boldsymbol{x} \\ \tilde{\boldsymbol{z}} \end{bmatrix}+\begin{bmatrix} \boldsymbol{B} \\ \boldsymbol{0} \end{bmatrix}\boldsymbol{v} \tag{19}$$

变成表示闭环系统的状态方程式. 显然, (19)式表明闭环系统的特征值由 ($\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K}$) 和 $\boldsymbol{F}$ 的特征值组成(参看 B-I-1 中系 1, 分块三角形矩阵行列式的性质).

证明完毕

P-I-5 可稳定性和可检测性

考虑常系数线性系统

$$\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{u},\ \boldsymbol{x}(0)=x_0,\ \boldsymbol{x}(t)\in R^n,\ \boldsymbol{u}(t)\in R^r \tag{1}$$

$$\boldsymbol{y}=\boldsymbol{C}\boldsymbol{x}+\boldsymbol{D}\boldsymbol{u},\ \boldsymbol{y}(t)\in R^n. \tag{2}$$

在 A-I-13 定理 4 中已经证明，当且仅当(1)式完全可控时，利用适当的状态反馈

$$\boldsymbol{u}=\boldsymbol{K}\boldsymbol{x}\quad \boldsymbol{K}\in R^{r\times n} \tag{3}$$

构成的闭环系统，其极点，即矩阵 $\hat{\boldsymbol{A}}\triangleq\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K}$ 的特征值(若为复根，则必须以共轭复根的形式成对地出现)可以自由配置．这里我们再稍微讨论一下可以自由选取 $\hat{\boldsymbol{A}}$ 的特征值的意义．对(1)式进行(3)式反馈得到的闭环系统是

$$\dot{\boldsymbol{x}}=\hat{\boldsymbol{A}}\boldsymbol{x},\ \boldsymbol{x}(0)=\boldsymbol{x}_0,\ \hat{\boldsymbol{A}}=\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K} \tag{4}$$

因此，其解可用下式给出(参看A-I-9)

$$\boldsymbol{x}(t)=e^{\hat{A}t}\boldsymbol{x}_0 \tag{5}$$

[引理 1] 当且仅当(1)式系统完全可控时，对于任意实数 M，选择某个适当的实数 $a>0$ 和矩阵 $\boldsymbol{K}$，对于任意 $\boldsymbol{x}_0\in R^n$，(5)式解 $\boldsymbol{x}(t)$ 有满足下式

$$\|\boldsymbol{x}(t)\|\leqslant a\|\boldsymbol{x}_0\|e^{Mt}\quad {}^\forall t\geqslant 0 \tag{6}$$

(注意)

若选成 $M<0$，则由(6)式得 $\|\boldsymbol{x}(t)\|\to 0\,(t\to\infty)$．对于任意 M(6)式成立意味着，可以使 $\|\boldsymbol{x}(t)\|$ 以任意指定的"衰减速度"趋近于 0．

(证明)由(5)式得

$$\|\boldsymbol{x}(t)\|\leqslant\|e^{\hat{A}t}\|\,\|\boldsymbol{x}_0\|. \tag{7}$$

利用 B-I-13 中(36)式将 $e^{\hat{A}t}$ 分解成 $\hat{\boldsymbol{A}}$ 的分量，得

$$e^{\hat{A}t}=\sum_{k=1}^{s}\sum_{l=0}^{d_k-1}\left[\frac{d^l}{d\lambda^l}e^{\lambda t}\right]_{\lambda=\lambda_k}\boldsymbol{E}_k^l=\sum_{k=1}^{s}\sum_{l=0}^{d_k-1}t^le^{\lambda_k t}\boldsymbol{E}_k^l \tag{8}$$

式中设 $\lambda_k(k=1,\cdots,s)$ 是 $\hat{\boldsymbol{A}}$ 的相异特征值，分别为 $\hat{\boldsymbol{A}}$ 的最小多项式的 d_k 重根．由(8)式得

$$\|e^{\hat{A}t}\|\leqslant\sum_{k=1}^{s}\sum_{l=0}^{d_k-1}|t^le^{\lambda_k t}|\,\|\boldsymbol{E}_k^l\|. \tag{9}$$

一般，对于任意 $\varepsilon>0$，存在某个 $a_{kl}(\varepsilon)>0$，使得

$$|t^le^{\lambda t}|=t^le^{\mathrm{Re}\,\lambda t}\leqslant a_{kl}\exp\{\mathrm{Re}\,\lambda+\varepsilon\}t\quad {}^\forall t\geqslant 0. \tag{10}$$

将该式代入(9)式，则存在某个常数 $a>0$，使得

$$\|e^{\hat{A}t}\|\leqslant a\exp\{\max_k\mathrm{Re}\,\lambda_k+\varepsilon\}t\quad {}^\forall t\geqslant 0. \tag{11}$$

因此，对于任意给定的 M，若选取使下式

$$\max_k\mathrm{Re}\,\lambda_k(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K})<\boldsymbol{M} \tag{12}$$

成立的 $\boldsymbol{K}$(若(1)式完全可控，则必然存在这样的 $\boldsymbol{K}$)，由(7)、(11)，(12)式，则(6)式成立.

反之，设(1)式$(\boldsymbol{A}, \boldsymbol{B})$不是完全可控，按照 A-I-11 定理 5 中用过的正则变换

$$\boldsymbol{x}=\boldsymbol{T}_c\boldsymbol{x}^c, \tag{13}$$

(1)式和下列系统代数等价

$$\dot{\boldsymbol{x}}^c=\boldsymbol{A}^c\boldsymbol{x}^c+\boldsymbol{B}^c\boldsymbol{u}$$

$$\boldsymbol{x}^c=\begin{bmatrix}\boldsymbol{x}_1^c\\ \boldsymbol{x}_2^c\end{bmatrix}\begin{matrix}\}\ q^c\\ \}\ n-q^c,\end{matrix}$$

$$\boldsymbol{A}^c\triangleq\boldsymbol{T}_c^{-1}\boldsymbol{A}\boldsymbol{T}_c=\begin{bmatrix}\boldsymbol{A}_{11}^c & \boldsymbol{A}_{12}^c\\ \underbrace{\boldsymbol{0}}_{q^c} & \underbrace{\boldsymbol{A}_{22}^c}_{n-q^c}\end{bmatrix}\begin{matrix}\}\ q^c\\ \}\ n-q^c\end{matrix}$$

$$\boldsymbol{B}^c\triangleq\boldsymbol{T}_c^{-1}\boldsymbol{B}=\begin{bmatrix}\boldsymbol{B}_1^c\\ \boldsymbol{0}\end{bmatrix}\begin{matrix}\}\ q^c\\ \}\ n-q^c\end{matrix} \tag{14}$$

这里$(\boldsymbol{A}_{11}^c, \boldsymbol{B}_1^c)$是可控的.

设反馈系数矩阵 $\boldsymbol{K}^c$ 为

$$\boldsymbol{K}^c\triangleq[\underbrace{\boldsymbol{K}_1^c}_{q^c}\ \underbrace{\boldsymbol{K}_2^c}_{n-q^c}], \tag{15}$$

则 $\boldsymbol{A}^c+\boldsymbol{B}^c\boldsymbol{K}^c$ 变成

$$\boldsymbol{A}^c+\boldsymbol{B}^c\boldsymbol{K}^c=\begin{bmatrix}\boldsymbol{A}_{11}^c+\boldsymbol{B}_1^c\boldsymbol{K}_1^c & \boldsymbol{A}_{12}^c+\boldsymbol{B}_1^c\boldsymbol{K}_2^c\\ \boldsymbol{0} & \boldsymbol{A}_{22}^c\end{bmatrix} \tag{16}$$

因$(\boldsymbol{A}_{11}^c, \boldsymbol{B}_1^c)$可控，则 $\boldsymbol{A}_{11}^c+\boldsymbol{B}_1^c\boldsymbol{K}_1^c$ 的特征值可任意选取，但是由(16)式显而易见，无论用什么样的 $\boldsymbol{K}^c$，$\boldsymbol{A}_{22}^c$ 的特征值都不变(参看 A-I-13)．因此，无论使用什么样的 $\boldsymbol{K}$，对于 $M<\max\mathrm{Re}\lambda(\boldsymbol{A}_{22}^c)$ 及任意 $\boldsymbol{x}_0\in R^n$，(6)式不可能成立.

证明完毕

关于系统的稳定性，大家知道有下面性质.

[引理 2] 对于任意 $\boldsymbol{x}_0\in R^n$，(5)式 $\boldsymbol{x}(t)$满足

$$\lim_{t\to\infty}\|\boldsymbol{x}(t)\|=0 \tag{17}$$

的充分必要条件是

$$\max\mathrm{Re}\,\lambda(\hat{\boldsymbol{A}})<0. \tag{18}$$

(证明) 充分性：设 $\max\mathrm{Re}\,\lambda(\hat{\boldsymbol{A}})=-\delta(\delta>0)$．若令 $\varepsilon\triangleq\delta/2$，则和(10)，(11)式同样

$$\|e^{\hat{A}t}\|\leqslant ae^{-\varepsilon t}\quad\forall t\geqslant 0. \tag{19}$$

因此，对于任意 $\boldsymbol{x}_0$(17)式成立.

必要性：例如，现假定 $\mathrm{Re}\,\lambda_1(\hat{\boldsymbol{A}})\geqslant 0$．因 $\boldsymbol{E}_1^0$ 不是零阵，则存在使 $\boldsymbol{E}_1^0\boldsymbol{x}_0\neq 0$ 成立的 $\boldsymbol{x}_0\in R^n$．因(8)式右边各项是独立的，则

$$\lim_{t\to\infty}e^{\lambda_1 t}\boldsymbol{E}_1^0\neq 0\Rightarrow\lim_{t\to\infty}\boldsymbol{x}(t)=\lim_{t\to\infty}e^{\hat{A}t}\boldsymbol{x}_0\neq 0.$$

因而，对于该 $\boldsymbol{x}_0$(17)式不成立.

证明完毕

考虑到该性质和(16)式，可见当且仅当 $\max \operatorname{Re} \lambda(\boldsymbol{A}_{22}^{c})<0$ 时，选择适当的 $\boldsymbol{K}$ 将会得到满足(17)式的(5)式解 $\boldsymbol{x}(t)$. 这里和完全可控的情况不同，“衰减速度”M 不能任意指定，即有 $M>\max \operatorname{Re} \lambda(\boldsymbol{A}_{22}^{c})$ 的限制. 虽然可能衰减得慢一些，但是无论如何解 $\boldsymbol{x}(t)$ 总会趋近原点. 针对着这一点，可以定义比完全可控性弱一些的如下性质.

[定义 1] 对系统(1)进行适当的反馈(3)，当闭环系统对于任意 $\boldsymbol{x}_0 \in R^n$ 有满足(17)式的解(5)时，系统(1)(或矩阵对$(\boldsymbol{A}, \boldsymbol{B})$)称为是可稳定的(Stabilizable).

由以上结果显然有下列性质.

[定理 1] 下面各条件均相互等价.

(i) 系统(1)是可稳定的.

(ii) 由某个矩阵 $\boldsymbol{K}$ 使得 $\max \operatorname{Re} \lambda(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K})<0$ 成立.

(iii) 进行(14)式的变换时，$\max \operatorname{Re} \lambda(\boldsymbol{A}_{22}^{c})<0$.

(iv) 任意的 $\boldsymbol{x}_0 \in R^n$ 至少具有下列中的一个性质:

(a) 是“可控的”的起始状态.

(b) $\lim t \to \infty\, e^{At}\boldsymbol{x}_0=\mathbf{0}$ 成立.

而且，作为比完全可观测性弱的有如下性质.

[定义 2] 当矩阵对$(\boldsymbol{A}^T, \boldsymbol{C}^T)$可稳定时，系统(1)，(2)(或矩阵对$(\boldsymbol{C}, \boldsymbol{A})$)称为是可检测的(detectable).

同时下列性质成立.

[定理 2] 下列各条件均相互等价.

(i) 系统(1)，(2)是可检测的.

(ii) 由某个矩阵 $\boldsymbol{G} \in R^{n\times m}$ 使得 $\max \operatorname{Re} \lambda(\boldsymbol{A}+\boldsymbol{G}\boldsymbol{C})<0$ 成立.

(iii) 当对系统(1)，(2)进行 A-I-11 定理 6 的变换时，$\max \operatorname{Re} \lambda(\boldsymbol{A}_{22}^{0})<0$.

(iv) 对于任意的 $\boldsymbol{x}_0 \in R^n$ 至少有下列中的一个性质成立:

(a) 是“可观测”的起始状态[1].

(b) $\lim\limits_{t\to\infty} e^{At}\boldsymbol{x}_0=\mathbf{0}$.

如 A-I-14 中所述，若系统(1)，(2)完全可观测，则利用 n 维或$(n-m)$维状态观测器可以产生状态 $\boldsymbol{x}$ 的估计值 $\boldsymbol{w}$，对于误差 $\boldsymbol{e} \triangleq \boldsymbol{x}-\boldsymbol{w}$ 可以给以任意的衰减率(参看 A-I-14 定理 1，2 的注意). 当系统(1)，(2)不是完全可观测而是可检测时，用 n 维或$(n-m)$维状态观测器估计状态不能任意指定衰减速度，但是无论如何可以使 $\boldsymbol{e}(t) \to \mathbf{0}$ $(t\to\infty)$ 成立.

(问题) 试证明上面性质.

1) “可观测”的意义可参照 A-I-11 中定理 7 的注意。

A-I-13′ 状 态 反 馈(续)

对于常系数线性系统

$$\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{u},\ \boldsymbol{x}(t)\in R^n,\ \boldsymbol{u}(t)\in R^r,\ \boldsymbol{A}\in R^{n\times n},\ \boldsymbol{B}\in R^{n\times n} \tag{1}$$

$$\boldsymbol{y}=\boldsymbol{C}\boldsymbol{x}+\boldsymbol{D}\boldsymbol{u},\ \boldsymbol{y}(t)\in R^m,\ \boldsymbol{C}\in R^{m\times n},\ \boldsymbol{D}\in R^{m\times r}, \tag{2}$$

控制输入 $\boldsymbol{u}$ 给成

$$\boldsymbol{u}=\boldsymbol{K}\boldsymbol{x}+\boldsymbol{v},\ \boldsymbol{v}(t)\in R^r,\ \boldsymbol{K}\in R^{r\times n} \tag{3}$$

称为状态反馈．它与系统的特性及控制有关，至于起到那些作用，在 A-I-13 中已讲过一些最基本的内容．此外，状态反馈还与许多问题有关，即使非常概括地介绍这些内容也需要相当的篇幅，而且这样也会离开本书的目的，这里仅指出它与那些问题有关，详细内容可参阅有关文献.

开始先假定(1)式起始条件 $\boldsymbol{x}(0)$ 为 $\boldsymbol{0}$，主要讨论输入输出关系.

模型适合问题

在(1)，(2)式系统中，从输入 $\boldsymbol{u}$ 到输出 $\boldsymbol{y}$ 的传递函数矩阵不用说可用

$$\boldsymbol{H}(s)=\boldsymbol{C}(s\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}+\boldsymbol{D} \tag{4}$$

给出．对(1)，(2)式系统进行状态空间的坐标变换

$$\boldsymbol{x}=\boldsymbol{T}\tilde{\boldsymbol{x}}\quad \boldsymbol{T}\in R^{n\times n},\ \det\boldsymbol{T}\neq 0, \tag{5}$$

传递函数矩阵仍然不变(参看 A-I-7，A-I-10)．但是进行输入坐标变换

$$\boldsymbol{u}=\boldsymbol{G}\tilde{\boldsymbol{u}},\ \boldsymbol{G}\in R^{r\times r},\ \det\boldsymbol{G}\neq 0 \tag{6}$$

及反馈变换((3)式)后，传递函数矩阵将发生变化.

因进行 $\boldsymbol{T}$ 变换并不影响传递函数矩阵，所以我们将 $\boldsymbol{T}$ 固定，令 $\boldsymbol{T}=\boldsymbol{I}_n$（这样并不失去普遍性），现在来考虑变换$(\boldsymbol{I}_n,\ \boldsymbol{G},\ \boldsymbol{K})$．对(1)，(2)式系统进行变换$(\boldsymbol{I}_n,\ \boldsymbol{G},\ \boldsymbol{K})$（参看 A-I-13），传递函数矩阵将由(4)式变成

$$\boldsymbol{H}(s;\boldsymbol{G},\boldsymbol{K})=(\boldsymbol{C}+\boldsymbol{D}\boldsymbol{K})(s\boldsymbol{I}-\boldsymbol{A}-\boldsymbol{B}\boldsymbol{K})^{-1}\boldsymbol{B}\boldsymbol{G}+\boldsymbol{D}\boldsymbol{G}. \tag{7}$$

这样，“进行变换$(\boldsymbol{I},\ \boldsymbol{G},\ \boldsymbol{K})$时传递函数矩阵如何变化”的问题就容易解决了．但是相反的问题，即“当给出系统(1)，(2)和传递函数矩阵 $\boldsymbol{H}_d(s)$ 时，是否存在使 $\boldsymbol{H}(s;\boldsymbol{G},\boldsymbol{K})$ 和 $\boldsymbol{H}_d(s)$ 一致的变换$(\boldsymbol{I},\ \boldsymbol{G},\ \boldsymbol{K})$，如果存在的话，怎样求?”．这个相反的问题称为模型适合(exact model matching)问题，它并不象前一个问题那么简单.

在模型适合问题中，所要求的变换$(\boldsymbol{I},\ \boldsymbol{G},\ \boldsymbol{K})$不一定存在．例如，在单输入单输出系统中，假定 $\boldsymbol{A}$, $\boldsymbol{b}$ 用可控典范形(参照 A-I-11，A-I-12)给出

$$\boldsymbol{A}=\begin{bmatrix} 0 & 1 & & \boldsymbol{0} \\ & \ddots & \ddots & \\ & \boldsymbol{0} & 0 & 1 \\ -a_n & -a_{n-1} & \cdots\ -a_2 & -a_1 \end{bmatrix}\quad \boldsymbol{b}=\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}. \tag{8}$$

设

$$\boldsymbol{c}=[c_1,\ c_2,\ \cdots,\ c_n]\quad d=0, \tag{9}$$

则通过简单的计算得

$$\boldsymbol{H}(s)=\frac{c_n s^{n-1}+c_{n-1}s^{n-2}+\cdots+c_2 s+c_1}{s^n+a_1 s^{n-1}+\cdots+a_{n-1}s+a_n}. \tag{10}$$

容易看到，进行反馈变换，(10)式的分母多项式将发生变化，而分子多项式不变．在这种情况下变换 $\boldsymbol{G}$ 只是乘以常数．也就是说，进行变换 $(\boldsymbol{I},\ \boldsymbol{G},\ \boldsymbol{K})$，(10)式传递函数的极点可以任意选择(参照 A-I-13)，但是零点完全不能改变．

那么，满足什么条件模型适合问题的解才存在呢？关于这个问题可参阅文献[102～104]等．

解耦控制问题

现在假定 $r=m$，由某个变换 $(\boldsymbol{I},\ \boldsymbol{G},\ \boldsymbol{K})$ 使 $\boldsymbol{H}(s;\boldsymbol{G},\ \boldsymbol{K})$ 变成正则对角矩阵．那么，v_i 对 $y_j(j\neq i)$ 的影响完全不出现．因此，操作 v_i 可以只控制 y_i，而对 $y_j(j\neq i)$ 没有影响，这样会带来很大的方便，称它为非干涉控制 (non-interacting control) 或解耦控制 (decoupling)．

系统(1)，(2)满足那些条件才可能进行解耦控制呢？换句话说，是否存在使 $\boldsymbol{H}(s;\boldsymbol{G},\ \boldsymbol{K})$ 变成正则对角矩阵的变换 $(\boldsymbol{I},\ \boldsymbol{G},\ \boldsymbol{K})$？这是解耦控制问题比较简单的情况．

在模型适合问题中是指定单一的传递函数矩阵 $\boldsymbol{H}_d(s)$，而这个解耦控制问题是指定正则对角矩阵类，考察是否存在变换矩阵 $(\boldsymbol{I},\ \boldsymbol{G},\ \boldsymbol{K})$，使 $\boldsymbol{H}(s,\ \boldsymbol{G},\ \boldsymbol{K})$ 变成该类中的元素．

由 B-I-1 中定理 2 得

$$\begin{aligned}
&\det(s\boldsymbol{I}-\boldsymbol{A}-\boldsymbol{BK})\det\boldsymbol{H}(s;\boldsymbol{G},\ \boldsymbol{K})\\
&=\det(s\boldsymbol{I}-\boldsymbol{A}-\boldsymbol{BK})\det\{(\boldsymbol{C}+\boldsymbol{DK})(s\boldsymbol{I}-\boldsymbol{A}-\boldsymbol{BK})^{-1}\boldsymbol{BG}+\boldsymbol{DG}\}\\
&=\det\begin{bmatrix} s\boldsymbol{I}-\boldsymbol{A}-\boldsymbol{BK} & \boldsymbol{BG}\\ -(\boldsymbol{C}+\boldsymbol{DK}) & \boldsymbol{DG}\end{bmatrix}\\
&=\det\begin{bmatrix} s\boldsymbol{I}-\boldsymbol{A} & \boldsymbol{B}\\ -\boldsymbol{C} & \boldsymbol{D}\end{bmatrix}\det\begin{bmatrix} \boldsymbol{I} & \boldsymbol{0}\\ -\boldsymbol{K} & \boldsymbol{G}\end{bmatrix}\\
&=\det(s\boldsymbol{I}-\boldsymbol{A})\det\{\boldsymbol{C}(s\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}+\boldsymbol{D}\}\det\boldsymbol{G}\\
&=\det(s\boldsymbol{I}-\boldsymbol{A})\det\boldsymbol{H}(s)\det\boldsymbol{G}
\end{aligned} \tag{11}$$

因其中 $(s\boldsymbol{I}-\boldsymbol{A}-\boldsymbol{BK})$、$(s\boldsymbol{I}-\boldsymbol{A})$，$\boldsymbol{G}$ 均为(有理函数体上的)正则矩阵，由(11)式，则当且仅当 $\boldsymbol{H}(s)$ 是正则矩阵时，$\boldsymbol{H}(s;\boldsymbol{G},\ \boldsymbol{K})$ 是正则矩阵．因此，上述解耦控制问题有解的必要条件是，$\det\boldsymbol{H}(s)\neq 0$，但它不是充分条件．在文献[105～107]中已求出上述解耦控制问题有解的充分必要条件．

以上讨论了用一个输入 v_i 控制一个输出 y_i，再稍为普遍一些，也可以考虑是否能用第 i 组输入 $\boldsymbol{v}_i$ 只控制第 i 组输出 $\boldsymbol{y}_i$，而不影响其它组输出 $\boldsymbol{y}_j(j\neq i)$ 的问题．在文献[105～109]中，对这种普遍性的问题也进行过很多研究，文献[110]出色地整理了这方面的发展．

逆系统

现在考虑 $r=m$，方阵 $\boldsymbol{H}(s)$ 是正则矩阵的情况. 若逆矩阵 $\boldsymbol{H}^{-1}(s)$ 是适宜的(参看A-I-10)，则存在传递函数矩阵等于 $\boldsymbol{H}^{-1}(s)$ 的常系数线性系统，称其为逆系统 (inverse system). 例如，在(2)式中 $\boldsymbol{D}$ 是正则矩阵的情况下，利用 B-I-1 中定理 3 可以得到

$$\begin{aligned}[\boldsymbol{H}(s)]^{-1} &= [\boldsymbol{C}(s\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}+\boldsymbol{D}]^{-1} \\ &= \boldsymbol{D}^{-1}-\boldsymbol{D}^{-1}\boldsymbol{C}(s\boldsymbol{I}-\boldsymbol{A}+\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C})^{-1}\boldsymbol{B}\boldsymbol{D}^{-1}. \end{aligned} \tag{12}$$

显然，该式右边变成常系数线性系统

$$\dot{\tilde{\boldsymbol{x}}} = (\boldsymbol{A}-\boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C})\tilde{\boldsymbol{x}}+\boldsymbol{B}\boldsymbol{D}^{-1}\tilde{\boldsymbol{u}} \tag{13}$$

$$\tilde{\boldsymbol{y}} = -\boldsymbol{D}^{-1}\boldsymbol{C}\tilde{\boldsymbol{x}}+\boldsymbol{D}^{-1}\tilde{\boldsymbol{u}} \tag{14}$$

的传递函数矩阵. 这里容易理解，(13)式相当于对(1)式进行变换 $(\boldsymbol{I},\ \boldsymbol{D}^{-1},\ -\boldsymbol{D}^{-1}\boldsymbol{C})$ 后得到的系统.

但是逆系统有什么意义呢？作出逆系统串联在原系统的前面(图 1)，由图 1 可见，若在逆系统的输入端加上 $\boldsymbol{w}(t)$，则原系统的输出与 $\boldsymbol{w}(t)$ 一致. 即，逆系统的存在和是否可以使系统输出 $\boldsymbol{y}(t)$，$0\leqslant t<\infty$ 与某个给定的时间函数 $\boldsymbol{w}(t)$ 一致的问题密切相关. 这里仍然假定 $\boldsymbol{x}(0)=\boldsymbol{0}$. 一般，在原系统的起始值 $\boldsymbol{x}(0)=\boldsymbol{x}_0\neq\boldsymbol{0}$ 的情况下，也象图 1 那样串联上逆系统，若对逆系统给以相同的起始值 $\tilde{\boldsymbol{x}}(0)=\boldsymbol{x}_0$，则当逆系统的输入端加上 $\boldsymbol{w}(t)$ 时，原系统的输出仍然与 $\boldsymbol{w}(t)$ 一致(试证明之). 因此，假定 $\boldsymbol{x}(0)=\boldsymbol{0}$ 并不失去普遍性.

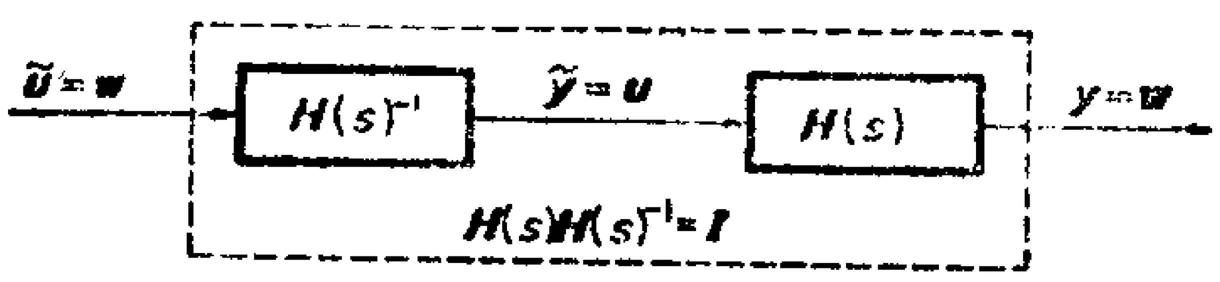


图 1

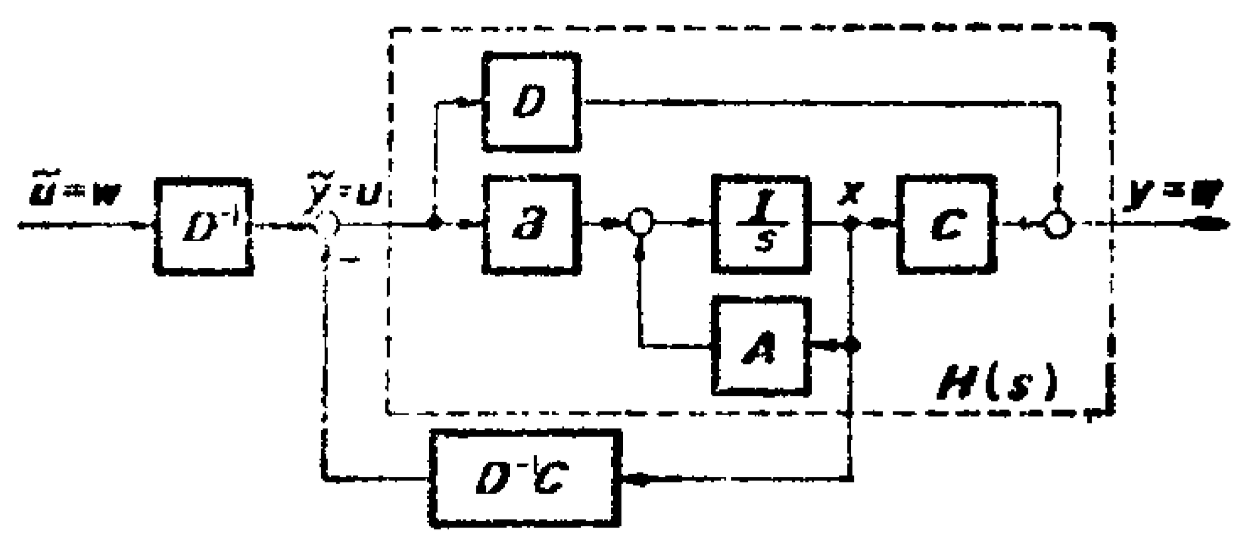


图 2

即使不象图 1 那样将 $\boldsymbol{H}^{-1}(s)$ 接到原系统上，例如象图 2 那样按照(13)，(14)式用模拟计算机实现，对给定系统仅仅进行变换 $(\boldsymbol{I},\ \boldsymbol{D}^{-1},\ -\boldsymbol{D}^{-1}\boldsymbol{C})$ 也可以得到和图 1 等价的输入输出关系. 从另外一个角度看，对于 $\boldsymbol{H}_d=\boldsymbol{I}$，这种情况下的模型适合问题可以解决.

一般在 $\boldsymbol{H}^{-1}(s)$ 不是适宜的情况下，即使其逆系统存在，也不能用标准形常系数线性系统实现，而必须采用微分器. 已经知道，在一定的条件下逆系统存在，它可以分成由微分器构成的部分和可以用标准形表示的部分，后者可以象(13)式那样写成对原系统进行某种变换 $(\boldsymbol{I},\ \boldsymbol{G},\ \boldsymbol{K})$ 的形式. 关于这方面的内容详见文献[111～115].

P-I-6 最优调节器问题

对于变系数线性系统

$$\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}+\boldsymbol{B}(t)\boldsymbol{u};\ \boldsymbol{x}(t_0)=\boldsymbol{x}_0$$
$$\boldsymbol{x}(t)\in R^n,\ \boldsymbol{u}(t)\in R^r,\ \boldsymbol{A}(t)\in R^{n\times n},\ \boldsymbol{B}(t)\in R^{n\times r} \tag{1}$$

考虑找出使泛函

$$J\triangleq\int_{t_0}^{t_f}\left\{\begin{bmatrix}\boldsymbol{x}(t)\\ \boldsymbol{u}(t)\end{bmatrix}^T\begin{bmatrix}\boldsymbol{W}_{11}(t) & \boldsymbol{W}_{12}(t)\\ \boldsymbol{W}_{12}^T(t) & \boldsymbol{W}_{22}(t)\end{bmatrix}\begin{bmatrix}\boldsymbol{x}(t)\\ \boldsymbol{u}(t)\end{bmatrix}\right\}dt+\boldsymbol{x}^T(t_f)\boldsymbol{P}_1\boldsymbol{x}(t_f) \tag{2}$$

为最小的控制输入 $\boldsymbol{u}_{[t_0,\,t_f]}$ 的最优控制问题. 其中假定, $\boldsymbol{A}(t)$, $\boldsymbol{B}(t)$, $\boldsymbol{W}_{11}(t)$, $\boldsymbol{W}_{12}(t)$, $\boldsymbol{W}_{22}(t)$ 都是以 t 的连续实函数为元素的有界矩阵, $\boldsymbol{W}_{22}(t)$ 对于所有的 $t\in[t_0,\ t_f]$ 都是对称正定且 $\boldsymbol{W}_{22}^{-1}(t)$ 有界, $\boldsymbol{W}_{11}(t)$ 及

$$\boldsymbol{W}(t)\triangleq\begin{bmatrix}\boldsymbol{W}_{11}(t) & \boldsymbol{W}_{12}(t)\\ \boldsymbol{W}_{12}^T(t) & \boldsymbol{W}_{22}(t)\end{bmatrix}\in R^{(n+r)\times(n+r)} \tag{3}$$

对于所有的 $t\in[t_0,\ t_f]$ 都是对称准正定的, $\boldsymbol{P}_1\in R^{n\times n}$ 是对称准正定的. 这个最优问题称为最优调节器问题 (optimal regulator problem).

这里要注意下列性质.

[引理 1] 当 $\boldsymbol{W}_{11}(t)$ 为对称准正定、$\boldsymbol{W}_{22}(t)$ 为对称正定时, (3) 式 $\boldsymbol{W}(t)$ 为对称准正定的充分必要条件是, $\boldsymbol{W}_{11}(t)-\boldsymbol{W}_{12}(t)\boldsymbol{W}_{22}^{-1}(t)\boldsymbol{W}_{12}^T(t)$ 为准正定的.

(证明) 因为显然

$$\boldsymbol{T}(t)\triangleq\begin{bmatrix}\boldsymbol{I}_n & \boldsymbol{0}_{n\times r}\\ -\boldsymbol{W}_{22}^{-1}(t)\boldsymbol{W}_{12}^T(t) & \boldsymbol{I}_r\end{bmatrix} \tag{4}$$

是正则矩阵, 若 (3) 式 $\boldsymbol{W}(t)$ 为准正定, 则

$$\boldsymbol{T}(t)^T\boldsymbol{W}(t)\boldsymbol{T}(t)=\begin{bmatrix}\boldsymbol{W}_{11}(t)-\boldsymbol{W}_{12}(t)\boldsymbol{W}_{22}(t)^{-1}\boldsymbol{W}_{12}(t)^T & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{W}_{22}(t)\end{bmatrix} \tag{5}$$

是准正定的. 因此, 在所有的 $t\in[t_0,\ t_f]$ 上 (5) 式右边矩阵的特征值都是非负的, 即 $\boldsymbol{W}_{11}(t)-\boldsymbol{W}_{12}(t)\boldsymbol{W}_{22}(t)^{-1}\boldsymbol{W}_{12}(t)^T$ 的特征值都是非负的. 因此, 该矩阵是准正定的. 考虑到 $\boldsymbol{W}_{22}(t)$ 的特征值始终都是正的, 将以上证明反推上去即可证明其逆定理.

证明完毕

若对 (1) 式系统进行下列状态反馈

$$\boldsymbol{u}=-[\boldsymbol{W}_{22}(t)^{-1}\boldsymbol{W}_{12}(t)^T]\boldsymbol{x}+\boldsymbol{v} \tag{6}$$

则

$$\dot{\boldsymbol{x}}=\{\boldsymbol{A}(t)-\boldsymbol{B}(t)\boldsymbol{W}_{22}(t)^{-1}\boldsymbol{W}_{12}(t)^T\}\boldsymbol{x}+\boldsymbol{B}(t)\boldsymbol{v},\ \ \boldsymbol{x}(t_0)=\boldsymbol{x}_0 \tag{7}$$

另外, 根据 (6) 式, 利用 (4) 式的 $\boldsymbol{T}(t)$ 可以写成

$$\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{u}\end{bmatrix}=\boldsymbol{T}(t)\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{v}\end{bmatrix}, \tag{8}$$

参照 (5) 式, 则 (2) 式的 J 可以改写成

$$J=\int_{t_0}^{t_f}\left\{\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{v}\end{bmatrix}^T\begin{bmatrix}\boldsymbol{W}_{11}(t)-\boldsymbol{W}_{12}(t)\boldsymbol{W}_{22}(t)^{-1}\boldsymbol{W}_{12}(t)^T & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{W}_{22}(t)\end{bmatrix}\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{v}\end{bmatrix}\right\}dt$$
$$+\boldsymbol{x}(t_f)^T\boldsymbol{P}_1\boldsymbol{x}(t_f). \tag{9}$$

现在对于(7)式，假定已经求出使(9)式泛函为最小的 $\boldsymbol{v}^*_{[t_0,t_f]}$，将其代入(7)式，设这时得到的解为 $\boldsymbol{x}^*_{[t_0,t_f]}$．将 $\boldsymbol{v}^*$，$\boldsymbol{x}^*$ 代入(6)式右边，设所得到的 $\boldsymbol{u}$ 为 $\boldsymbol{u}^*$，则 $\boldsymbol{u}^*_{[t_0,t_f]}$ 是对于(1)式使(2)式 J 为最小的最优控制输入，$\boldsymbol{x}^*_{[t_0,t_f]}$ 和将 $\boldsymbol{u}^*_{[t_0,t_f]}$ 代入(1)式得到的最佳轨线一致.

由上面的结果可以看到，为了解决由(1)，(2)式规定的一般最优调节器问题．实际上解出如下限定的最优调节器问题是充分必要的.

“对系统

$$\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}+\boldsymbol{B}(t)\boldsymbol{u},\quad \boldsymbol{x}(t_0)=\boldsymbol{x}_0 \tag{10}$$

求使泛函

$$J=\int_{t_0}^{t_f}\{\boldsymbol{x}(t)^T\boldsymbol{Q}(t)\boldsymbol{x}(t)+\boldsymbol{u}(t)^T\boldsymbol{R}(t)\boldsymbol{u}(t)\}dt+\boldsymbol{x}(t_f)^T\boldsymbol{P}_1\boldsymbol{x}(t_f) \tag{11}$$

为最小的控制输入 $\boldsymbol{u}_{[t_0,t_f]}$．其中假定 $\boldsymbol{A}(t)$，$\boldsymbol{B}(t)$，$\boldsymbol{Q}(t)$，$\boldsymbol{R}(t)$ 都是以 t 的连续实函数为元素的有界矩阵，在所有的 $t\in[t_0,\ t_f]$ 上，$\boldsymbol{Q}(t)\in R^{n\times n}$ 是对称准正定、$\boldsymbol{R}(t)\in R^{r\times r}$ 是对称正定且 $\boldsymbol{R}^{-1}_{(t)}$ 有界，同时 $\boldsymbol{P}_1\in R^{n\times n}$ 是对称准正定的．”

关于这个最优调节器问题，许多文献[67, 116~120]都有详细说明，故以下仅摘录一些重要的结果.

［定理 1］“最优调节器问题中的最优输入 $\boldsymbol{u}^0(t)$ 可以由下列状态反馈

$$\boldsymbol{u}^0(t)=-\boldsymbol{K}^0(t)\boldsymbol{x}^0(t) \tag{12}$$
$$\boldsymbol{K}^0(t)=\boldsymbol{R}(t)^{-1}\boldsymbol{B}(t)^T\boldsymbol{P}(t) \tag{13}$$

给出．式中对称准正定矩阵 $\boldsymbol{P}$ 是黎卡提(Riccati)*微分方程式

$$-\dot{\boldsymbol{P}}(t)=\boldsymbol{Q}(t)-\boldsymbol{P}(t)\boldsymbol{B}(t)\boldsymbol{R}(t)^{-1}\boldsymbol{B}(t)^T\boldsymbol{P}(t)+\boldsymbol{P}(t)\boldsymbol{A}(t)+\boldsymbol{A}(t)^T\boldsymbol{P}(t) \tag{14}$$

满足下列终端条件

$$\boldsymbol{P}(t_f)=\boldsymbol{P}_1 \tag{15}$$

的解．而且，用该最优控制时 J 的最小值变成

$$J_{\min}\triangleq\int_{t_0}^{t_f}\{\boldsymbol{x}^0(t)^T\boldsymbol{Q}(t)\boldsymbol{x}^0(t)+\boldsymbol{u}^0(t)^T\boldsymbol{R}(t)\boldsymbol{u}^0(t)\}dt\boldsymbol{x}(t_f)^T\boldsymbol{P}_1\boldsymbol{x}(t_f)$$
$$=\boldsymbol{x}(t_0)^T\boldsymbol{P}(t_0)\boldsymbol{x}(t_0) \tag{16}$$

因 $\boldsymbol{Q}(t)$ 是准正定的，则可利用某个矩阵 $\boldsymbol{C}(t)$ 将其分解成 $\boldsymbol{Q}(t)=\boldsymbol{C}(t)^T\boldsymbol{C}(t)$，$\boldsymbol{C}(t)$ 也是以 t 的连续实函数为元素的矩阵.

［定理 2］“若［$\boldsymbol{A}(t)$，$\boldsymbol{B}(t)$ 完全可控，当 $t_f\to\infty$ 时，则满足终端条件 $\boldsymbol{P}(t_f)=\boldsymbol{0}$ 的(14)式的解收敛于准正定矩阵 $\overline{\boldsymbol{P}}(t)$，$\overline{\boldsymbol{P}}(t)$ 是满足(14)式的解.

而且，若［$\boldsymbol{A}(t)$，$\boldsymbol{B}(t)$］一致完全可控，［$\boldsymbol{A}(t)$，$\boldsymbol{C}(t)$］一致完全可观测，设 $\boldsymbol{P}_1$ 为任意准正定矩阵，当 $t_f\to\infty$ 时，则满足 $\boldsymbol{P}(t_f)=\boldsymbol{P}_1$ 的(14)式的解收敛于 $\overline{\boldsymbol{P}}(t)$。

对于任意 $\boldsymbol{P}_1\geqslant\boldsymbol{0}$，状态反馈控制

$$\boldsymbol{u}(t)=-\boldsymbol{R}(t)^{-1}\boldsymbol{B}(t)^T\overline{\boldsymbol{P}}(t)\boldsymbol{x}(t) \tag{17}$$

---

* 括弧内原文系译者注。

使

$$\lim_{t_f\to\infty}\int_{t_0}^{t_f}\{\boldsymbol{x}(t)^T\boldsymbol{Q}(t)\boldsymbol{x}(t)+\boldsymbol{u}(t)^T\boldsymbol{R}(t)\boldsymbol{u}(t)\}dt+\boldsymbol{x}(t_f)^T\boldsymbol{P}_1\boldsymbol{x}(t_f) \tag{18}$$

为最小. 其最小值可以用 $\boldsymbol{x}(t_0)^T\overline{\boldsymbol{P}}(t_0)\boldsymbol{x}(t_0)$ 给出."

现在我们考虑常系数的情况，假定 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{Q}$, $\boldsymbol{R}$ 均为不依存于 t 的实矩阵.

[定理 3] "设 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{Q}$, $\boldsymbol{R}$ 均为实常数矩阵. 这时有下列性质:

(a) 根据 A-I-11 中定理 6 的变换，将 $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ 改写成下列形式

$$\dot{\boldsymbol{x}}^0=\begin{bmatrix}\boldsymbol{A}_{11}^0 & \boldsymbol{0}\\ \boldsymbol{A}_{21}^0 & \boldsymbol{A}_{22}^0\end{bmatrix}\boldsymbol{x}^0+\begin{bmatrix}\boldsymbol{B}_1^0\\ \boldsymbol{B}_2^0\end{bmatrix}\boldsymbol{u}$$

$\boldsymbol{C}\boldsymbol{x}=[\boldsymbol{C}_1^0\ \ \boldsymbol{0}]\boldsymbol{x}^0$, 矩阵对 $(\boldsymbol{C}_1^0, \boldsymbol{A}_{11}^0)$ 完全可观测,

在这里，当且仅当矩阵对 $(\boldsymbol{A}_{11}^0, \boldsymbol{B}_1^0)$ 可稳定时，以 $\boldsymbol{P}_1=\boldsymbol{0}$ 为起始条件的黎卡提方程式的解，在 $t_f\to\infty$ 时收敛于准正定常数矩阵 $\overline{\boldsymbol{P}}$.

(b) 若 $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ 可稳定且可检测，则以任意 $\boldsymbol{P}_1\geqslant\boldsymbol{0}$ 为起始条件的黎卡提方程式的解，在 $t_f\to\infty$ 时收敛于同一个准正定的 $\overline{\boldsymbol{P}}$.

(c) 若 $\overline{\boldsymbol{P}}$ 存在，则它是代数方程式

$$\boldsymbol{Q}-\boldsymbol{P}\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^T\boldsymbol{P}+\boldsymbol{A}^T\boldsymbol{P}+\boldsymbol{P}\boldsymbol{A}=\boldsymbol{0} \tag{19}$$

的准正定解. 若 $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ 可稳定且可检测，则 $\overline{\boldsymbol{P}}$ 是(19)式唯一的准正定解.

(d) 设 $\overline{\boldsymbol{P}}$ 存在，$\overline{\boldsymbol{P}}$ 正定的充分必要条件是 $(\boldsymbol{A}, \boldsymbol{C})$ 完全可观测.

(e) 设 $\overline{\boldsymbol{P}}$ 存在,

$$\max \operatorname{Re}\lambda(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^T\overline{\boldsymbol{P}})<0 \tag{20}$$

成立的充分必要条件是 $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ 可稳定且可检测.

(f) 若 $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ 可稳定且可检测，则对于任意 $\boldsymbol{P}_1\geqslant 0$, 状态反馈控制

$$\boldsymbol{u}(t)=-\boldsymbol{R}^{-1}\boldsymbol{B}^T\overline{\boldsymbol{P}}\boldsymbol{x}(t) \tag{21}$$

使

$$\lim_{t_f\to\infty}\int_{t_0}^{t_f}\{\boldsymbol{x}(t)^T\boldsymbol{Q}\boldsymbol{x}(t)+\boldsymbol{u}(t)^T\boldsymbol{R}\boldsymbol{u}(t)\}dt+\boldsymbol{x}(t_f)^T\boldsymbol{P}_1\boldsymbol{x}(t_f) \tag{22}$$

为最小，该最小值可用 $\boldsymbol{x}(t_0)^T\overline{\boldsymbol{P}}\boldsymbol{x}(t_0)$ 给出."

A-I-13″ 状 态 反 馈(续)

在 P-I-6 中已经证明过，一般对于变系数线性系统，使状态及控制输入的二次性能指标为最小的最优控制输入，为变系数状态反馈．而且，特别是对于常系数线性系统

$$\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{u},\ \boldsymbol{x}(t)\in R^{n},\ \boldsymbol{u}(t)\in R^{r}, \tag{1}$$

使二次性能指标

$$J=\int_{t_0}^{\infty}\{\boldsymbol{x}(t)^{T}\boldsymbol{Q}\boldsymbol{x}(t)+\boldsymbol{u}(t)^{T}\boldsymbol{R}\boldsymbol{u}(t)\}dt \tag{2}$$

为最小的最优控制输入变成常系数状态反馈

$$\boldsymbol{u}^{0}(t)=\boldsymbol{K}\boldsymbol{x}(t) \tag{3}$$

$$\boldsymbol{K}=-\boldsymbol{R}^{-1}\boldsymbol{B}^{T}\boldsymbol{P}, \tag{4}$$

其中矩阵 $\boldsymbol{P}$ 是代数方程式

$$\boldsymbol{Q}-\boldsymbol{P}\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^{T}\boldsymbol{P}+\boldsymbol{A}^{T}\boldsymbol{P}+\boldsymbol{P}\boldsymbol{A}=\boldsymbol{0} \tag{5}$$

的准正定解．

但是反过来，假定用某个任意矩阵 $\boldsymbol{K}\in R^{r\times n}$ 对系统(1)进行(3)式的状态反馈，控制输入 $\boldsymbol{u}^{0}$ 是否能成为使某种性能指标为最小的最优控制呢?

这个问题可以说是最优控制的逆问题，文献[121～126]等曾讨论过．

A-I-15　变系数线性系统中的状态反馈

前章介绍了关于常系数线性系统中状态反馈的各种结果，这里大概讲一下这些结果在变系数情况下变得如何.

对于变系数线性系统

$$\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}+\boldsymbol{B}(t)\boldsymbol{u}\quad \boldsymbol{x}(t)\in R^n,\ \boldsymbol{u}(t)\in R^r,\ \boldsymbol{A}(t)\in R^{n\times n},\ \boldsymbol{B}(t)\in R^{n\times r} \tag{1}$$

$$\boldsymbol{y}=\boldsymbol{C}(t)\boldsymbol{x}\qquad \boldsymbol{y}(t)\in R^m,\ \boldsymbol{C}(t)\in R^{m\times n}, \tag{2}$$

考虑下列三种变换:

状态空间的坐标变换

$$\tilde{\boldsymbol{x}}=\boldsymbol{T}(t)\boldsymbol{x}\quad \boldsymbol{T}(t)\in R^{n\times n} \tag{3}$$

输入空间的坐标变换

$$\boldsymbol{u}=\boldsymbol{G}(t)\tilde{\boldsymbol{u}}\quad \boldsymbol{G}(t)\in R^{r\times r} \tag{4}$$

以及状态变量反馈

$$\boldsymbol{u}=\boldsymbol{K}(t)\boldsymbol{x}+\boldsymbol{v}\quad \boldsymbol{v}(t)\in R^r,\ \boldsymbol{K}(t)\in R^{r\times n}{}^{*} \tag{5}$$

这里假定 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{T}$, $\boldsymbol{G}$, $\boldsymbol{K}$ 都是以 t 的连续实函数为元素的矩阵，$\boldsymbol{T}(t)$, $\boldsymbol{G}(t)$ 在所有的 $t\in R$ 上都是正则的.

仿照 A-I-13，将按照(5)，(4)，(3)顺序所进行的变换略记为$(\boldsymbol{T},\ \boldsymbol{G},\ \boldsymbol{K})_t$. 由满足上述条件的矩阵组成的变换$(\boldsymbol{T},\ \boldsymbol{G},\ \boldsymbol{K})_t$全体的集合设其为 $\mathscr{G}_t$. $\mathscr{G}_t$ 等价性也和 A-I-13 中同样定义.

可控性

［定理 1］　对于 $\mathscr{G}$ 等价系统，完全可控性保存.

（证明）　和 A-I-13 中定理 1 的“证明方法 1”一样.

可控系统的典范形

假定在时间坐标轴的某个区间 J 上 $\boldsymbol{A}$, $\boldsymbol{B}$ 的元素无限次连续可微（即是 C^∞ 级）. 对于变换矩阵 $\boldsymbol{T}$, $\boldsymbol{G}$, $\boldsymbol{K}$ 的元素再加上在 J 上是 C^∞ 级的条件. 将满足这个条件的变换略记为$(\boldsymbol{T},\ \boldsymbol{G},\ \boldsymbol{K})_t^\infty$，设其全体的集合为 $\mathscr{G}_t^\infty$.

在区间 J 上定义(参照 A-I-8)

$$\boldsymbol{P}_0(t)\triangleq\boldsymbol{B}(t) \tag{6}$$

$$\begin{aligned}&\boldsymbol{P}_j(t)\triangleq-\boldsymbol{A}(t)\boldsymbol{P}_{j-1}(t)+\dot{\boldsymbol{P}}_{j-1}(t),\ j\geqslant1,\\&\boldsymbol{Q}_k(t)\triangleq[\boldsymbol{P}_0(t),\ \boldsymbol{P}_1(t),\ \cdots,\ \boldsymbol{P}_{k-1}(t)],\\&k=1,\ 2,\ \cdots,\ n-1,\end{aligned} \tag{7}$$

* 原文误印成 $\boldsymbol{K}(t)\boldsymbol{R}^{r\times n}$.——译者注

而且设

$$\left.\begin{aligned} & q_k(t) \triangleq \operatorname{rank} \boldsymbol{Q}_k(t) \\ & r_k(t) \triangleq q_{k+1}(t) - q_k(t) \\ & r_0(t) \triangleq q_1(t). \end{aligned}\right\} \quad k=1, \cdots, n-1 \tag{8}$$

于是有如下结果[93](证明省略,仅列出其结果).

[定理 2]　$(\boldsymbol{A}(t), \boldsymbol{B}(t))$在$J$上与某个可控常系数线性系统$(\boldsymbol{A}, \boldsymbol{B})\mathscr{G}_t^\infty$等价的充分必要条件是

$r_0(t), r_1(t), \cdots, r_{n-1}(t)$在$J$上是一定的,

$r_0(t)+r_1(t)+\cdots+r_{n-1}(t)=n$ 成立.

而且, 若$r_0(t)=r$, $\forall t \in J$, 则$(\boldsymbol{A}(t), \boldsymbol{B}(t))$与A-I-13中(25)式的典范形$(\boldsymbol{A}_0, \boldsymbol{B}_0)\mathscr{G}_t^\infty$等价.

状态反馈形成的闭环系统特征值的配置

在A-I-13及A-I-14中, 对于常系数线性系统$\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{u}$, $\boldsymbol{y}=\boldsymbol{C}\boldsymbol{x}+\boldsymbol{D}\boldsymbol{u}$所讲过的状态反馈具有的效果中, 最基本的是闭环系统有特征值可以配置的性质 (A-I-13中定理4). 而且,当状态不能直接测定,而要通过观测器间接进行状态反馈时, 由于含有观测器的复合闭环系统的特征值可以分离成原闭环系统的特征值和观测器的特征值. 则状态反馈增益的设定和观测器的设计也可以分别考虑.

下面将要讲一下,对于变系数线性系统,这个特征值配置的性质以怎样的形式成立.

在变系数系统中, 因为特征值(或传递函数的极点)的概念不具有象常系数系统那样的意义,所以不能按照配置其特征值的观点处理闭环系统的动态特性. 因此,我们要考察比特征值更为基本的量——闭环系统的状态转移矩阵. 为了适当设定反馈增益$\boldsymbol{K}(\cdot)$使闭环系统的状态转移矩阵(或同样,零输入响应特性)可控, 由常系数系统类推, 关于控制对象$\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}+\boldsymbol{B}(t)\boldsymbol{u}$的那一些可控性必须成立. 因此, 将特征值的配置问题作如下考虑,即"若控制对象具有某种程度的可控性,是否可以将由状态反馈$\boldsymbol{u}=\boldsymbol{K}(t)\boldsymbol{x}+\boldsymbol{v}$形成的闭环系统的零输入响应的过渡特性设定在指定的形式? 若使用观测器, 情况又如何呢?"

对于变系数系统

$$\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}+\boldsymbol{B}(t)\boldsymbol{u},\ \boldsymbol{A}(t)\in R^{n\times n},\ \boldsymbol{B}(t)\in R^{n\times r} \tag{9a}$$

$$\boldsymbol{y}=\boldsymbol{C}(t)\boldsymbol{x}+\boldsymbol{D}(t)\boldsymbol{u},\ \boldsymbol{C}(t)\in \boldsymbol{R}^{m\times n},\ \boldsymbol{D}(t)\in R^{m\times r}, \tag{9b}$$

若用状态反馈

$$\boldsymbol{u}=\boldsymbol{K}(t)\boldsymbol{x}+\boldsymbol{v} \tag{5}$$

则闭环系统状态方程式及输出方程式分别变成

$$\dot{\boldsymbol{x}}=(\boldsymbol{A}(t)+\boldsymbol{B}(t)\boldsymbol{K}(t))\boldsymbol{x}+\boldsymbol{B}(t)\boldsymbol{v}, \tag{10a}$$

$$\boldsymbol{y}=(\boldsymbol{C}(t)+\boldsymbol{D}(t)\boldsymbol{K}(t))\boldsymbol{x}+\boldsymbol{D}(t)\boldsymbol{v}. \tag{10b}$$

在这里考虑, 闭环状态方程式(10a)的零输入响应$\dot{\boldsymbol{x}}=(\boldsymbol{A}(t)+\boldsymbol{B}(t)\boldsymbol{K}(t))\boldsymbol{x}$的过渡特性按指数函数设定,使之满足下列关系

$$a\|\boldsymbol{x}(t_0)\|e^{m(t-t_0)}\leqslant\|\boldsymbol{x}(t)\|\leqslant b\|\boldsymbol{x}(t_0)\|e^{M(t-t_0)}\quad \forall t\geqslant t_0,\ \forall t_0\in \boldsymbol{R}. \tag{11}$$

(11)式对于任意$\boldsymbol{x}(t_0)$均成立与$(\boldsymbol{A}(t)+\boldsymbol{B}(t)\boldsymbol{K}(t))$的转移矩阵$\tilde{\boldsymbol{\Phi}}$满足

$$\|\widetilde{\boldsymbol{\Phi}}(t,\ t_0)\| \leqslant b e^{M(t-t_0)}, \qquad \forall t \geqslant t_0 \tag{12a}$$

$$\|\widetilde{\boldsymbol{\Phi}}^{-1}(t,\ t_0)\| \leqslant a^{-1} e^{-m(t-t_0)}, \quad \forall t \geqslant t_0 \tag{12b}$$

等价.

（问题 1） 试证明(11)式和(12)式等价.

"提示：
$$\max_{\boldsymbol{x}(t_0)\neq 0} \|\boldsymbol{x}(t)\| / \|\boldsymbol{x}(t_0)\| = \|\widetilde{\boldsymbol{\Phi}}(t,\ t_0)\|,$$
$$\min_{\boldsymbol{x}(t_0)\neq 0} \|\boldsymbol{x}(t)\| / \|\boldsymbol{x}(t_0)\| = \|\widetilde{\boldsymbol{\Phi}}^{-1}(t,\ t_0)\|^{-1}.\text{''}$$

那么.对于上述问题的基本结果可以给出如下.

［定理 3］ "若控制对象 $\dot{\boldsymbol{x}} = \boldsymbol{A}(t)\boldsymbol{x} + \boldsymbol{B}(t)\boldsymbol{u}$ 一致完全可控，则对于任意实数 $M,\ m$ $(M \geqslant m)$，关于闭环系统 $\dot{\boldsymbol{x}} = (\boldsymbol{A}(t) + \boldsymbol{B}(t)\boldsymbol{K}(t))\boldsymbol{x}$ 的任意解 $\boldsymbol{x}(t)$，可以选择出使(11)式成立的常数 $a,\ b\,(b \geqslant a > 0)$ 和反馈矩阵 $\boldsymbol{K}(\cdot)$."

定理 3 可以看成是与常系数系统特征值配置相对应的变系数情况下的结果．该定理的详细证明请参考文献[127]，这里只是扼要地说明一下大致证明过程．为了说明方便，分成下列三个阶段：(i)若一致完全可控性的第 1 个条件(13a)式成立，对于任意 M，则从上面限制成 $\|\boldsymbol{x}(t)\| \leqslant b\|\boldsymbol{x}(t_0)\| e^{M(t-t_0)}$；(ii)若第 2 个条件 (13b) 式成立，则意味着对于任意 m 从下面限制成 $a\|\boldsymbol{x}(t_0)\| e^{m(t-t_0)} \leqslant \|\boldsymbol{x}(t)\|$；(iii)而且若第 1，2 个条件同时成立（即一致完全可控)，当 $M,\ m\,(M \geqslant m)$ 同时给定时，可以象(11)式那样从上下指定 $\|\boldsymbol{x}(t)\|$.

$\dot{\boldsymbol{x}} = \boldsymbol{A}(t)\boldsymbol{x} + \boldsymbol{B}(t)\boldsymbol{u}$ 一致完全可控在物理意义上意味着，将 $\boldsymbol{x}(t)$ 由任意起始状态 $\boldsymbol{x}(t_0)$ 移到原点 $\boldsymbol{0}$ 时，再由原点到达任意目标 $\boldsymbol{x}_T$ 时所需要的最小输入能量，与起始时间无关，而分别与 $\|\boldsymbol{x}(t_0)\|^2$，$\|\boldsymbol{x}_T\|^2$ 成比例.

若用式子表示，设可控性矩阵为

$$\boldsymbol{W}_c(t_1,\ t_2) \triangleq \int_{t_1}^{t_2} \boldsymbol{\Phi}(t_1,\ \tau)\boldsymbol{B}(\tau)\boldsymbol{B}^T(\tau)\boldsymbol{\Phi}^T(t_1,\ \tau)d\tau,$$

则有某些正实数 $\sigma,\ \alpha_0,\ \alpha_1,\ \beta_0,\ \beta_1$，使得

$$\boldsymbol{0} < \alpha_0\boldsymbol{I} \leqslant \boldsymbol{W}_c(t,\ t+\sigma) \leqslant \alpha_1\boldsymbol{I}, \quad \forall t \in R \tag{13a}$$

$$\boldsymbol{0} < \beta_0\boldsymbol{I} \leqslant \boldsymbol{\Phi}(t+\sigma,\ t)\boldsymbol{W}_c(t,\ t+\sigma)\boldsymbol{\Phi}^T(t+\sigma,\ t) \leqslant \beta_1\boldsymbol{I}, \quad \forall t \in R \tag{13b}$$

成立(A-I-8 中定义 4)．式中 $\boldsymbol{\Phi}$ 是 $\boldsymbol{A}(\cdot)$ 的状态转移矩阵．在(13b)式中将 $t+\sigma$ 改写成 t，因 $\boldsymbol{W}_c$ 可按上述给出，则(13b)式可换成下列等价条件

$$\boldsymbol{0} < \beta_0\boldsymbol{I} \leqslant \boldsymbol{Y}_c(t-\sigma,\ t) \leqslant \beta_1\boldsymbol{I}, \quad \forall t \in R \tag{13b$'$}$$

其中

$$\begin{aligned}\boldsymbol{Y}_c(t-\sigma,\ t) &\triangleq \boldsymbol{\Phi}(t,\ t-\sigma)\boldsymbol{W}_c(t-\sigma,\ t)\boldsymbol{\Phi}^T(t,\ t-\sigma)\\ &= \int_{t-\sigma}^{t} \boldsymbol{\Phi}(t,\ \tau)\boldsymbol{B}(\tau)\boldsymbol{B}^T(\tau)\boldsymbol{\Phi}(t,\ \tau)d\tau.\end{aligned} \tag{14}$$

(i) (13a)$\Rightarrow\|\boldsymbol{x}(t)\| \leqslant b\|\boldsymbol{x}(t_0)\| e^{M(t-t_0)}$ 的证明

考虑将$\{\boldsymbol{A}(\cdot),\ \boldsymbol{B}(\cdot)\}$的可控性矩阵 $\boldsymbol{W}_c$ 作若干修正后的矩阵

$$\widetilde{\boldsymbol{W}}_c(t,\ t+\sigma) \triangleq \int_t^{t+\sigma} \boldsymbol{\Phi}(t,\ \tau)\boldsymbol{B}(\tau)\boldsymbol{B}^T(\tau)\boldsymbol{\Phi}^T(t,\ \tau)e^{-2M(t-\tau)}d\tau, \tag{15}$$

它相当于$\{\boldsymbol{A}(\cdot) - M\boldsymbol{I},\ \boldsymbol{B}(\cdot)\}$的可控性矩阵．因(13a)式成立，则对于 $\widetilde{\boldsymbol{W}}_c$，

$$\boldsymbol{0} < \alpha_0 e^{-2|M|\sigma}\boldsymbol{I} \leqslant \widetilde{\boldsymbol{W}}_c(t,\ t+\sigma) \leqslant \alpha_1 e^{2|M|\sigma}\boldsymbol{I} \tag{16}$$

成立．因 $\widetilde{\boldsymbol{W}}_c(t,\ t+\sigma)$ 是正则的，则其逆矩阵存在，用它作为反馈矩阵，试确定

$$\boldsymbol{K}(t)=-\frac{1}{2}\boldsymbol{B}^T(t)\widetilde{\boldsymbol{W}}_c^{-1}(t,\ t+\sigma). \tag{17}$$

在这里考虑二次型

$$V(\boldsymbol{x},\ t)=\boldsymbol{x}^T\widetilde{\boldsymbol{W}}_c^{-1}(t,\ t+\sigma)\boldsymbol{x}. \tag{18}$$

若考虑到(16)式,因该二次型满足

$$\alpha_1^{-1}e^{-2|M|\sigma}\|\boldsymbol{x}\|^2\leqslant V(\boldsymbol{x},\ t)\leqslant\alpha_0^{-1}e^{2|M|\sigma}\|\boldsymbol{x}\|^2,\quad \forall t\in R,\ \forall \boldsymbol{x}\in R^n, \tag{19}$$

则它是正值二次型.(19)式中使用的是欧几里得范数.因 R^n 中所有的范数均等价,则使用欧几里得范数证明已足够.以后在本章中向量及矩阵的范数都使用欧几里得范数.

那么,对于反馈采用(17)式 $\boldsymbol{K}(\cdot)$ 的闭环系统 $\dot{\boldsymbol{x}}=(\boldsymbol{A}(t)+\boldsymbol{B}(t)\boldsymbol{K}(t))\boldsymbol{x}$,我们来沿其解计算一下 $V(\boldsymbol{x}(t),\ t)$ 的微分.考虑到

$$\begin{aligned}\frac{d}{dt}\widetilde{\boldsymbol{W}}_c(t,\ t+\sigma)=&-\boldsymbol{B}(t)\boldsymbol{B}^T(t)+\widetilde{\boldsymbol{\Phi}}(t,\ t+\sigma)\boldsymbol{B}(t+\sigma)\boldsymbol{B}^T(t+\sigma)\\&\times\boldsymbol{\Phi}^T(t,\ t+\sigma)e^{2M\sigma}+(\boldsymbol{A}(t)-M\boldsymbol{I})\widetilde{\boldsymbol{W}}_c(t,\ t+\sigma)\\&+\widetilde{\boldsymbol{W}}_c(t,\ t+\sigma)(\boldsymbol{A}(t)-M\boldsymbol{I}),\end{aligned}$$

则

$$\begin{aligned}\dot{V}(\boldsymbol{x}(t),\ t)=&\boldsymbol{x}(t)[2M\widetilde{\boldsymbol{W}}_c^{-1}(t,\ t+\sigma)-\widetilde{\boldsymbol{W}}_c^{-1}(t,\ t+\sigma)\boldsymbol{\Phi}(t,\ t+\sigma)\boldsymbol{B}(t+\sigma)\\&\times\boldsymbol{B}^T(t+\sigma)\boldsymbol{\Phi}^T(t,\ t+\sigma)e^{2M\sigma}\widetilde{\boldsymbol{W}}_c^{-1}(t,\ t+\sigma)]\boldsymbol{x}(t)\\&\leqslant 2MV(\boldsymbol{x}(t),\ t),\quad \forall t\in R.\end{aligned}$$

令 $\dot{V}(\boldsymbol{x}(t),\ t)=2MV(\boldsymbol{x}(t),\ t)-\alpha(t)$, $\alpha(t)\geqslant 0$,将 $\alpha(\cdot)$ 视为输入,对 $V(\boldsymbol{x}(t),\ t)$ 解该微分方程式,得 $V(\boldsymbol{x}(t),\ t)\leqslant V(\boldsymbol{x}(t_0),\ t_0)e^{2M(t-t_0)}$, $\forall t\geqslant t_0$, $\forall t_0\in R$.而且考虑到(19)式,则

$$\|\boldsymbol{x}(t)\|^2\leqslant\alpha_1\alpha_0^{-1}e^{4|M|\sigma}\|\boldsymbol{x}(t_0)\|^2e^{2M(t-t_0)}, \tag{20}$$

于是证明了所要的结果.

(ii) $(13\text{b})'\Rightarrow a\|\boldsymbol{x}(t_0)\|e^{m(t-t_0)}\leqslant\|\boldsymbol{x}(t)\|$ 的证明

该证明几乎和(i)一样.设对应于$\{\boldsymbol{A}(\cdot)-m\boldsymbol{I},\ \boldsymbol{B}(\cdot)\}$的 $\boldsymbol{Y}_c$ 矩阵为 $\widetilde{\boldsymbol{Y}}_c$,即令

$$\widetilde{\boldsymbol{Y}}_c(t-\sigma,\ t)\triangleq\int_{t-\sigma}^{t}\boldsymbol{\Phi}(t,\ \tau)\boldsymbol{B}(\tau)\boldsymbol{B}^T(\tau)\boldsymbol{\Phi}(t,\ \tau)e^{-2m(t-\tau)}d\tau, \tag{21}$$

反馈矩阵按

$$\boldsymbol{K}(t)=\frac{1}{2}\boldsymbol{B}^T(t)\widetilde{\boldsymbol{Y}}_c^{-1}(t-\sigma,\ t) \tag{22}$$

确定,沿闭环系统的解计算正值二次型 $V(\boldsymbol{x}(t),\ t)=\boldsymbol{x}^T\widetilde{\boldsymbol{Y}}_c^{-1}(t-\sigma,\ t)\boldsymbol{x}$ 对时间的微分,得

$$\dot{V}(\boldsymbol{x}(t),\ t)\geqslant 2mV(\boldsymbol{x}(t),\ t). \tag{23}$$

将(23)式写成 $\dot{V}(\boldsymbol{x}(t),\ t)=2mV(\boldsymbol{x}(t),\ t)+\alpha(t)$, $\alpha(t)\geqslant 0$, $\alpha(\cdot)$视为输入,解该微分方程式得

$$V(\boldsymbol{x}(t),\ t)\geqslant V(\boldsymbol{x}(t_0),\ t_0)e^{2m(t-t_0)}. \tag{24}$$

因由(24)式可得

$$\|\boldsymbol{x}(t)\|^2\geqslant\beta_0\beta_1^{-1}e^{-4|m|\sigma}\|\boldsymbol{x}(t_0)\|^2e^{2m(t-t_0)},\quad \forall t\geqslant t_0,$$

设 $a\triangleq\sqrt{\beta_0/\beta_1}e^{-2|m|\sigma}$,则所要的结果成立.

(iii) 因(13a),$(13\text{b})\Rightarrow a\|\boldsymbol{x}(t_0)\|e^{m(t-t_0)}\leqslant\|\boldsymbol{x}(t)\|\leqslant b\|\boldsymbol{x}(t_0)\|e^{M(t-t_0)}$ 的证明非常复杂,故省略[127].这里我们感兴趣的是,在上下时间常数设定得相等的情况下($M=m$),也可以作出实现它的适当的反馈增益.如下面所述,在使用观测器的系统中,在 $M>m$ 的范

围内必须指定 M 和 m.

(注意)

在定理 3 中, $\{\boldsymbol{A}(\cdot),\ \boldsymbol{B}(\cdot)\}$ 的一致完全可控性是闭环系统的零输入响应可以象(11)式那样从上下用指数函数加以限制的充分条件. 若为有界系统($\|\boldsymbol{A}(t)\|$ 和 $\|\boldsymbol{B}(t)\|$ 有界), 也可以证明实际上这是充分必要条件[127]. 这个结果与"常系数系统 $\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{u}$ 中闭环系统特征值配置的可能性和 $\{\boldsymbol{A},\ \boldsymbol{B}\}$ 的可控性等价"相对应.

使用观测器的状态反馈

和常系数的情况一样, 对于变系数线性系统(9)也可以考虑 $\boldsymbol{K}(t)\boldsymbol{x}$-观测器及 $(n-m)$ 维状态观测器. 但是这里不打算讨论观测器的构成问题, 主要考虑通过观测器进行状态反馈时闭环系统总的特性. 因此, 我们只考察最基本的 n 维状态观测器.

对于系统(9), 当定义 n 维系统

$$\dot{\boldsymbol{z}}=\boldsymbol{F}(t)\boldsymbol{z}+\boldsymbol{G}(t)\boldsymbol{u}+\boldsymbol{H}(t)\boldsymbol{y} \tag{25}$$

的系数矩阵 $\boldsymbol{F}(t)=\boldsymbol{A}(t)-\boldsymbol{H}(t)\boldsymbol{C}(t),\ \boldsymbol{G}(t)=\boldsymbol{B}(t)-\boldsymbol{H}(t)\boldsymbol{D}(t)$ 时, 因

$$\dot{\boldsymbol{e}}=[\boldsymbol{A}(t)-\boldsymbol{H}(t)\boldsymbol{C}(t)]\boldsymbol{e},\quad \boldsymbol{e}\triangleq\boldsymbol{x}-\boldsymbol{z} \tag{26}$$

成立, 若对于任意起始误差 $\boldsymbol{e}(t_0)$ (26)式的解为 $\boldsymbol{e}(t)\to\boldsymbol{0},\ t\to\infty$, 则系统(25)是(9)式的 n 维状态观测器.

也就是说, 可以这样来设定 $\boldsymbol{H}(\cdot)$, 使(26)式具有上述那样的性质. 我们知道, 根据定理 3, 若对于 $\{\boldsymbol{A}(\cdot),\ \boldsymbol{B}(\cdot)\}$ (13a) 式成立, 则可以适当地选择 $\boldsymbol{K}(\cdot)$, 对 $\dot{\boldsymbol{x}}=[\boldsymbol{A}(t)+\boldsymbol{B}(t)\boldsymbol{K}(t)]\boldsymbol{x}$ 的解给出任意的衰减率. 但是因为(26)式中 $\boldsymbol{H}(t)$ 的位置在 $\boldsymbol{C}(t)$ 的前面, 所以这个结果不能直接用于(26)式. 在常系数系统中, 这个区别不是本质的. 其原因是, $\dot{\boldsymbol{x}}=[\boldsymbol{A}-\boldsymbol{H}\boldsymbol{C}]\boldsymbol{x}$ 的解的衰减性由其系数矩阵 $[\boldsymbol{A}-\boldsymbol{H}\boldsymbol{C}]$ 的特征值决定, $[\boldsymbol{A}-\boldsymbol{H}\boldsymbol{C}]$ 的特征值与 $[\boldsymbol{A}^T-\boldsymbol{C}^T\boldsymbol{H}^T]$ 的特征值一致. 对于变系数系统, 例如在 $\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}$ 和 $\dot{\tilde{\boldsymbol{x}}}=\boldsymbol{A}^T(t)\tilde{\boldsymbol{x}}$ 的解之间, 如下面例子所示, 没有常系数系统那种关系.

[例]

$$\boldsymbol{A}(t)=\begin{bmatrix}-1 & e^{2t}\\ 0 & -4\end{bmatrix}\quad 及\quad \boldsymbol{A}^T(t)=\begin{bmatrix}-1 & 0\\ e^{2t} & -4\end{bmatrix}$$

的转移矩阵分别为

$$\boldsymbol{\Phi}(t,\ t_0)=\begin{bmatrix}e^{-(t-t_0)} & e^{-(2t-4t_0)}+e^{-(t-3t_0)}\\ 0 & e^{-4(t-t_0)}\end{bmatrix}$$

$$\tilde{\boldsymbol{\Phi}}(t,\ t_0)=\begin{bmatrix}e^{-(t-t_0)} & 0\\ \dfrac{1}{5}e^{t+t_0}-\dfrac{1}{5}e^{-(4t-6t_0)} & e^{-4(t-t_0)}\end{bmatrix}.$$

因此, $\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}$ 的解全部为 $\boldsymbol{x}(t)\to\boldsymbol{0}$, 而对于 $\dot{\tilde{\boldsymbol{x}}}=\boldsymbol{A}^T(t)\tilde{\boldsymbol{x}}$, 只要 $\tilde{\boldsymbol{x}}_1(t_0)\neq0$, $\|\tilde{\boldsymbol{x}}(t)\|$ 就是发散的.

为了利用[定理 3]的结果, 我们重新考虑系统(9)的对偶系统

$$\dot{\tilde{\boldsymbol{x}}}=-\boldsymbol{A}^T(t)\tilde{\boldsymbol{x}}-\boldsymbol{C}^T(t)\tilde{\boldsymbol{u}} \tag{27a}$$

$$\tilde{\boldsymbol{y}}=\boldsymbol{B}^T(t)\tilde{\boldsymbol{x}}+\boldsymbol{D}^T(t)\tilde{\boldsymbol{u}}. \tag{27b}$$

若对系统(27)进行状态反馈 $\tilde{\boldsymbol{u}}=-\boldsymbol{H}^T(t)\tilde{\boldsymbol{x}}$, 则闭环系统的状态方程式为

$$\dot{\tilde{\boldsymbol{x}}}=[-\boldsymbol{A}^T(t)+\boldsymbol{C}^T(t)\boldsymbol{H}^T(t)]\tilde{\boldsymbol{x}}. \tag{28}$$

这刚好是系统(26)的伴随系统．但是伴随系统的“不稳定化”对应于原系统的“稳定化”．即，对于任意正实数 m，选择某个 $\boldsymbol{H}^T(\cdot)$ 和常数 $a>0$，设系统(28)的解为

$$\|\tilde{\boldsymbol{x}}(t)\|\geqslant a\|\tilde{\boldsymbol{x}}(t_0)\|e^{m(t-t_0)},\ \forall t\geqslant t_0,\ \forall t_0\in \boldsymbol{R}, \tag{29}$$

则用相同 $\boldsymbol{H}(\cdot)$ 的系统(26)的解满足

$$\|\boldsymbol{e}(t)\|\leqslant a^{-1}\|\boldsymbol{e}(t_0)\|e^{-m|t-t_0|},\ \forall t\geqslant t_0,\ \forall t_0\in \boldsymbol{R}. \tag{30}$$

设系统(26)的转移矩阵为 $\boldsymbol{\Phi}(t,\ t_0)$，则系统(28)的转移矩阵为 $\boldsymbol{\Phi}^T(t_0,\ t)$，$\|\boldsymbol{\Phi}^T(t_0,\ t)\|=\|\boldsymbol{\Phi}(t_0,\ t)\|$ 成立(注意，使用的是欧几里得范数)，由此很容易推导出上述结果．

但是根据[定理 3]，对于(27a)式，使(29)式成立的反馈 $\boldsymbol{H}^T(\cdot)$ 存在的充分条件是，关于 $\{-\boldsymbol{A}^T(\cdot),\ -\boldsymbol{C}^T(\cdot)\}$ (13b)式成立．即，设 $\boldsymbol{A}(t)$ 的转移矩阵为 $\boldsymbol{\Phi}(t,\ t_0)$ 时

$$0<\beta_0\boldsymbol{I}\leqslant\boldsymbol{\Phi}^T(t,\ t+\sigma)\boldsymbol{W}_0(t,\ t+\sigma)\boldsymbol{\Phi}(t,\ t+\sigma)\leqslant\beta_1\boldsymbol{I} \tag{31}$$

式中

$$\boldsymbol{W}_0(t,\ t+\sigma)=\int_t^{t+\sigma}\boldsymbol{\Phi}^T(\tau,\ t)\boldsymbol{C}^T(\tau)\boldsymbol{C}(\tau)\boldsymbol{\Phi}(\tau,\ t)d\tau.$$

如 A-I-8 中所述，当 $\boldsymbol{W}_0(t,\ t+\sigma)$ 按上式确定时，系统(9)一致完全可观测就是(31)式和

$$\boldsymbol{0}<\alpha_0\boldsymbol{I}\leqslant\boldsymbol{W}_0(t,\ t+\sigma)\leqslant\alpha_1\boldsymbol{I} \tag{32}$$

同时成立．结果可以看到，系统(9)一致完全可观测的条件之一(31)式成立时，选择适当的 $\boldsymbol{H}(\cdot)$，对观测器(25)的误差可以给出任意的指数函数衰减速度．

最后我们再来研究一下，当系统(1)用 n 维状态观测器(25)进行状态反馈时(如图 3 所示)复合闭环系统的零输入响应变得如何．

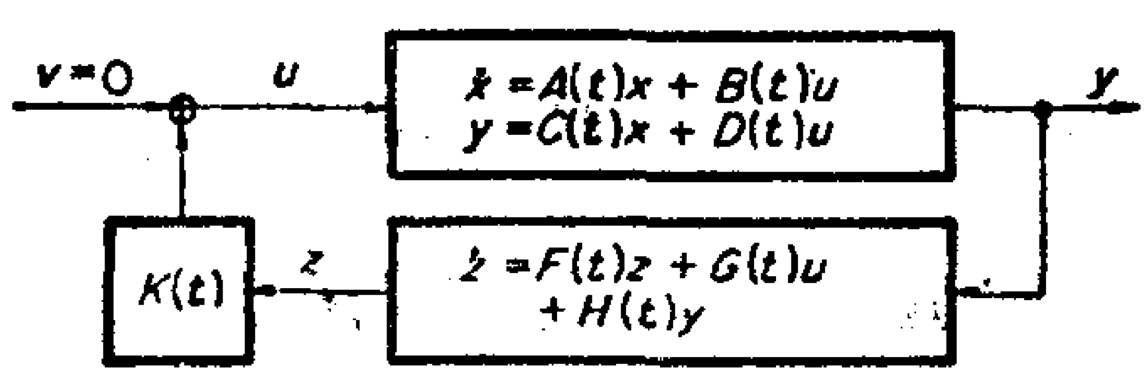


图 3　使用观测器的闭环系统

首先，假定 $\boldsymbol{v}=\boldsymbol{0}$，对于 $\boldsymbol{x}$ 和 $\boldsymbol{z}$ 的闭环系统状态方程式为

$$\begin{bmatrix}\dot{\boldsymbol{x}}\\ \dot{\boldsymbol{z}}\end{bmatrix}=\begin{bmatrix}\boldsymbol{A}(t) & \boldsymbol{B}(t)\boldsymbol{K}(t)\\ \boldsymbol{H}(t)\boldsymbol{C}(t) & \boldsymbol{A}(t)-\boldsymbol{H}(t)\boldsymbol{C}(t)+\boldsymbol{B}(t)\boldsymbol{K}(t)\end{bmatrix}\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{z}\end{bmatrix}. \tag{33}$$

利用 $\boldsymbol{e}\triangleq\boldsymbol{x}-\boldsymbol{z}$，上式进一步变成

$$\begin{bmatrix}\dot{\boldsymbol{x}}\\ \dot{\boldsymbol{e}}\end{bmatrix}=\begin{bmatrix}\boldsymbol{A}(t)+\boldsymbol{B}(t)\boldsymbol{K}(t) & -\boldsymbol{B}(t)\boldsymbol{K}(t)\\ \boldsymbol{0} & \boldsymbol{A}(t)-\boldsymbol{H}(t)\boldsymbol{C}(t)\end{bmatrix}\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{e}\end{bmatrix}^*. \tag{34}$$

由(34)式显然，若是常系数系统，因闭环系统的特征值由 $(\boldsymbol{A}+\boldsymbol{BK})$ 和 $(\boldsymbol{A}-\boldsymbol{HC})$ 的特征值构成，只考虑特征值则 $\boldsymbol{K}$ 和 $\boldsymbol{H}$ 完全可以独立地设定．在变系数系统中，因涉及到闭环系统的转移矩阵，故不能完全独立地确定．

现在假定，闭环系统(34)的转移矩阵为 $\boldsymbol{\Phi}_T$，与 $\boldsymbol{A}(t)+\boldsymbol{B}(t)\boldsymbol{K}(t)$ 及 $\boldsymbol{A}(t)-\boldsymbol{H}(t)\boldsymbol{C}(t)$ 对应的转移矩阵分别为 $\boldsymbol{\Phi}_c$，$\boldsymbol{\Phi}_e$．利用 $\boldsymbol{\Phi}_c$，$\boldsymbol{\Phi}_e$ 可将 $\boldsymbol{\Phi}_T$ 表示成(试证明之)

* 原文误印成 $\begin{bmatrix}\dot{\boldsymbol{x}}\\ \dot{\boldsymbol{e}}\end{bmatrix}=\begin{bmatrix}\boldsymbol{A}(t)+\boldsymbol{B}(t)\boldsymbol{K}(t) & -(t)\boldsymbol{K}(t)\\ \boldsymbol{0} & \boldsymbol{A}(t)-\boldsymbol{H}(t)\boldsymbol{C}(t)\end{bmatrix}\begin{bmatrix}\boldsymbol{x}\\ \boldsymbol{e}\end{bmatrix}$. ——译者注

$$\boldsymbol{\Phi}_T(t,\ t_0)=\begin{bmatrix}\boldsymbol{\Phi}_c(t,\ t_0) & -\int_{t_0}^{t}\boldsymbol{\Phi}_c(t,\ \tau)\boldsymbol{B}(\tau)\boldsymbol{K}(\tau)\boldsymbol{\Phi}_e(\tau,\ t_0)d\tau\\ \mathbf{0} & \boldsymbol{\Phi}_e(t,\ t_0)\end{bmatrix}. \tag{35}$$

在这里，考虑对闭环系统(27)的零输入响应的过渡过程给出

$$\left\|\begin{matrix}\boldsymbol{x}(t)\\ \boldsymbol{e}(t)\end{matrix}\right\|\leqslant b\left\|\begin{matrix}\boldsymbol{x}(t_0)\\ \boldsymbol{e}(t_0)\end{matrix}\right\|e^{-M|t-t_0|},\quad \forall t\geqslant t_0,\quad \forall t_0\in R \tag{36}$$

的衰减特性，式中 M 是任意正实数. 如问题 1 中所述，(36)式和

$$\|\boldsymbol{\Phi}_T(t,\ t_0)\|\leqslant be^{-M(t-t_0)},\quad \forall t\geqslant t_0,\quad \forall t_0\in R \tag{37}$$

等价，由(35)式可以得出[1)]

$$\|\boldsymbol{\Phi}_T(t,\ t_0)\|\leqslant\|\boldsymbol{\Phi}_c(t,\ t_0)\|+\|\boldsymbol{\Phi}_e(t,\ t_0)\|+\left\|\int_{t_0}^{t}\boldsymbol{\Phi}_c(t,\ \tau)\boldsymbol{B}(\tau)\boldsymbol{K}(\tau)\boldsymbol{\Phi}_e(\tau,\ t_0)d\tau\right\| \tag{38}$$

由此显然，虽然假定 $\|\boldsymbol{\Phi}_c\|$ 和 $\|\boldsymbol{\Phi}_e\|$ 各用 $\exp[-M(t-t_0)]$ 加以限制，但是由于(38)式第三项的影响，不能说(37)式就成立. 因此现在假定，将 $\boldsymbol{K}(\cdot)$ 和 $\boldsymbol{H}(\cdot)$ 选择成各只具有 $\exp[-\varepsilon(t-t_0)]$ 的裕度，

$$\begin{aligned}&\|\boldsymbol{\Phi}_c(t,\ t_0)\|\leqslant b_1e^{-(M+\varepsilon)(t-t_0)}\\ &\|\boldsymbol{\Phi}_e(t,\ t_0)\|\leqslant b_2e^{-(M+\varepsilon)(t-t_0)}\end{aligned} \tag{39}$$

成立. 式中 ε 是任意正实数. 由以上讲过的内容可见，对于系统(9)，若一致完全可控性，一致完全可观测性同时成立，则上述的 $\boldsymbol{K}(\cdot)$，$\boldsymbol{H}(\cdot)$ 可以分别用(17)，(22)式的形式构成[2)]. 现在我们由(39)式来证明，(38)式右边第 3 项可以用 $\exp[-M(t-t_0)]$ 限制.

首先

$$\begin{aligned}&\left\|\int_{t_0}^{t}\boldsymbol{\Phi}_c(t,\ \tau)\boldsymbol{B}(\tau)\boldsymbol{K}(\tau)\boldsymbol{\Phi}_e(\tau,\ t_0)\,d\tau\right\|\\ &\leqslant b_1b_2e^{-(M+\varepsilon)(t-t_0)}\int_{t_0}^{t}\|\boldsymbol{B}(\tau)\boldsymbol{K}(\tau)\|d\tau.\end{aligned} \tag{40}$$

在这里，因 $\boldsymbol{K}(\cdot)$ 由(17)式确定，$\boldsymbol{K}(t)=-\frac{1}{2}\boldsymbol{B}^T(t)\widehat{\boldsymbol{W}}_c^{-1}(t-\sigma,\ t)$，由(16)式，$\left\|\frac{1}{2}\widehat{\boldsymbol{W}}_c^{-1}(t-\sigma,\ t)\right\|\leqslant k<\infty$，则(40)式右边变成

$$\leqslant b_1b_2ke^{-(M+\varepsilon)(t-t_0)}\int_{t_0}^{t}\|\boldsymbol{B}(\tau)\boldsymbol{B}^T(\tau)\|d\tau. \tag{41}$$

现在，若假定系统(9)一致完全可控，则有适当的正实数 C_1，C_2 存在，使得

$$\int_{t_0}^{t}\|\boldsymbol{B}(\tau)\boldsymbol{B}^T(\tau)\|d\tau\leqslant C_1+C_2(t-t_0),\quad \forall t\geqslant t_0,\quad \forall t_0\in R \tag{42}$$

成立(其证明在本章附录中给出)，利用该结果，则(41)式最后变成

$$\begin{aligned}&\leqslant b_1b_2ke^{-M(t-t_0)}e^{-\varepsilon(t-t_0)}[C_1+C_2(t-t_0)]\\ &\leqslant b_3e^{-M(t-t_0)},\quad b_3\triangleq b_1b_2k(C_1+C_2\varepsilon^{-1}e^{-1}).\end{aligned} \tag{43}$$

设 $b\triangleq b_1+b_2+b_3$，由(39)，(43)式可以证明，所要求的结果(37)式成立. 将以上归纳如下.

［定理 4］ 若系统(9)一致完全可控且一致完全可观测，则对于任意给定的 $M>0$，

---

1) 将 $\left\|\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \mathbf{0} & \boldsymbol{C}\end{bmatrix}\right\|^2=\max\limits_{\|\boldsymbol{x}\|=1}\left\|\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \mathbf{0} & \boldsymbol{C}\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}\right\|^2\leqslant(\|\boldsymbol{A}\|+\|\boldsymbol{B}\|)^2+\|\boldsymbol{C}\|^2\leqslant(\|\boldsymbol{A}\|+\|\boldsymbol{B}\|+\|\boldsymbol{C}\|)^2$ 的结果用于(35)式.

2) 在这个阶段，若(13a)，(31)式成立就足够了.

在使用观测器(25)的闭环系统(34)中,有使

$$\left\|\begin{matrix}\boldsymbol{x}(t)\\ \boldsymbol{e}(t)\end{matrix}\right\|\leqslant b\left\|\begin{matrix}\boldsymbol{x}(t_0)\\ \boldsymbol{e}(t_0)\end{matrix}\right\|e^{-M(t-t_0)},\quad \forall t\geqslant t_0,\quad \forall t_0\in \boldsymbol{R}$$

成立的正实数 b 和变系数反馈矩阵 $\boldsymbol{K}(\cdot)$ 及观测器矩阵 $\boldsymbol{H}(\cdot)$ 存在.

(注意)

在定理 4 中考虑用指数函数作为闭环系统零输入响应的衰减上限，而实际上可以证明[127],在同样条件下(即系统(9)一致完全可控且一致完全可观测)，当给出任意实数 M, m(而 $M>m$)时,和定理 3 一样，可以用 $b\exp[M(t-t_0)]$ 和 $a\exp[m(t-t_0)]$ 给出衰减上、下限．但是要注意，象定理 3 那样，可以直接得到状态反馈时，也可以令 $M=m$，但是在通过观测器的反馈中不能令 $M=m$．如下面例子所示,考虑到常系数的情况,就立即明白其道理了．因为作为定理 2 的特殊情况也含常系数系统,若这时不能令 $M=m$,当然一般情况下就更不可能了．

(例)　设纯量常系数系统 $\dot{x}=ax+bu$, $y=cx+du$ 的 1 维状态观测器为 $\dot{z}=(a-hc)z+(b-hd)u+hy$．使用状态反馈 $u=kz$ 时,闭环系统为

$$\begin{bmatrix}\dot{x}\\ \dot{e}\end{bmatrix}=\begin{bmatrix}a+bk & -bk\\ 0 & a-hc\end{bmatrix}\begin{bmatrix}x\\ e\end{bmatrix}. \tag{44}$$

假定其中反馈常数 k 和观测器常数 h 分别按满足

$$a_1\left\|\begin{matrix}x(0)\\ e(0)\end{matrix}\right\|e^{Mt}\leqslant\left\|\begin{matrix}x(t)\\ e(t)\end{matrix}\right\|\leqslant b_1\left\|\begin{matrix}x(0)\\ e(0)\end{matrix}\right\|e^{Mt} \tag{45}$$

的条件确定,则 M 是任意实数．但是必须使 $a+bk=a-hc=M$．即假定 $k=b^{-1}(M-a)$, $h=c^{-1}(a-M)$．根据(35)式,这时闭环系统(44)的转移矩阵为

$$\boldsymbol{\Phi}_T(t,\ t_0)=\begin{bmatrix}e^{M(t-t_0)} & -bk(t-t_0)e^{M(t-t_0)}\\ 0 & e^{M(t-t_0)}\end{bmatrix},$$

由于其中包含 te^{Mt} 形的元素,所以不存在使 $\|\boldsymbol{\Phi}(t,\ t_0)\|\leqslant b_1\exp[M(t-t_0)]$ 成立的常数 b_1, 即不能选出使得(45)式成立的 k, h, a_1, b_1.

附录

$\int_{t_0}^{t}\|\boldsymbol{B}(\tau)\boldsymbol{B}^T(\tau)\|d\tau\leqslant C_1+C_2(t-t_0)$ 的证明

假定$\{\boldsymbol{A}(\cdot),\ \boldsymbol{B}(\cdot)\}$一致完全可控,即设(13)式成立.

$$\begin{aligned}&\int_{t_0}^{t_0+\sigma}\|\boldsymbol{B}(\tau)\boldsymbol{B}^T(\tau)\|d\tau\\ &=\int_{t_0}^{t_0+\sigma}\|\boldsymbol{\Phi}(\tau,\ t_0)\boldsymbol{\Phi}(t_0,\ \tau)\boldsymbol{B}(\tau)\boldsymbol{B}^T(\tau)\boldsymbol{\Phi}^T(t_0,\ \tau)\boldsymbol{\Phi}^T(\tau,\ t_0)\|d\tau\\ &\leqslant\sup_{t_0\leqslant\tau\leqslant t_0+\sigma}\|\boldsymbol{\Phi}(\tau,\ t_0)\|^2\int_{t_0}^{t_0+\sigma}\|\boldsymbol{\Phi}(t_0,\ \tau)\boldsymbol{B}(\tau)\boldsymbol{B}^T(\tau)\boldsymbol{\Phi}^T(t_0,\ \tau)\|d\tau.\end{aligned}$$

在这里，若$\{\boldsymbol{A}(\cdot),\ \boldsymbol{B}(\cdot)\}$一致完全可控，利用卡尔曼的结果[67]，即有使$\|\boldsymbol{\Phi}(t,\ s)\|\leqslant\gamma(|t-s|)$, $\forall t,\ s\in \boldsymbol{R}$ 成立的函数 $\gamma(\cdot)$存在,则可以证明

$$
\begin{aligned}
&\int_{t_0}^{t_0+\sigma}\|\boldsymbol{B}(\tau)\boldsymbol{B}^T(\tau)\|d\tau\\
&\leqslant \sup_{0\leqslant\tau\leqslant\sigma}\gamma^2(\tau)\int_{t_0}^{t_0+\tau}\|\boldsymbol{\Phi}(t_0,\ \tau)\boldsymbol{B}(\tau)\boldsymbol{B}^T(\tau)\boldsymbol{\Phi}^T(t_0,\ \tau)\|d\tau\\
&\leqslant \sup_{0\leqslant\tau\leqslant\sigma}\gamma^2(\tau)\int_{t_0}^{t_0+\sigma}\mathrm{tr}[\boldsymbol{\Phi}(t_0,\ \tau)\boldsymbol{B}(\tau)\boldsymbol{B}^T(\tau)\boldsymbol{\Phi}^T(t_0,\ \tau)]d\tau\\
&= \sup_{0\leqslant\tau\leqslant\sigma}\gamma^2(\tau)\,\mathrm{tr}\int_{t_0}^{t_0+\sigma}\boldsymbol{\Phi}(t_0,\ \tau)\boldsymbol{B}(\tau)\boldsymbol{B}^T(\tau)\boldsymbol{\Phi}^T(t_0,\ \tau)d\tau\\
&= \sup_{0\leqslant\tau\leqslant\sigma}\gamma^2(\tau)\,\mathrm{tr}\,\boldsymbol{W}_c(t_0,\ t+\sigma)\\
&\leqslant \sup_{0\leqslant\tau\leqslant\sigma}\gamma^2(\tau)n\alpha_2\\
&\triangleq C_1.
\end{aligned}
$$

因此，一般

$$
\int_{t_0}^{t}\|\boldsymbol{B}(\tau)\boldsymbol{B}^T(\tau)\|d\tau\leqslant C_1\left\{1+\frac{1}{\sigma}(t-t_0)\right\}\triangleq C_1+C_2(t-t_0).
$$

在以上证明中，使用了关于准正定对称矩阵 $\boldsymbol{A}(\tau)\in R^{n\times n}$ 的 $\|\boldsymbol{A}(\tau)\|=\lambda_{\max}(\boldsymbol{A}(\tau))\leqslant\mathrm{tr}\boldsymbol{A}(\tau)$ 的性质(B-I-11 中系 23)及

$$
\int_{t_0}^{t_0+\sigma}\mathrm{tr}\boldsymbol{A}(\tau)\,d\tau=\int_{t_0}^{t_0+\sigma}[a_{11}(\tau)+\cdots+a_{nn}(\tau)]d\tau=\mathrm{tr}\int_{t_0}^{t_0+\sigma}\boldsymbol{A}(\tau)d\tau
$$

的结果.

B-I-14 广义逆矩阵[128~130]

本章将讨论在系统理论中广泛应用的广义逆矩阵的基本内容. 从字面上理解,所谓广义逆矩阵(generalized inverse)是普通意义上逆矩阵的推广,其必要性可以说是起始于求线性方程组的解,这时也包含其系数矩阵不是正则矩阵的情况.

若 $\boldsymbol{A}$ 为正则方阵,则线性方程组

$$\boldsymbol{Ax}=\boldsymbol{y},\ \boldsymbol{A}\in C^{m\times n},\ \boldsymbol{y}\in C^{m},\ \boldsymbol{x}\in C^{n} \tag{1}$$

的解 $\boldsymbol{x}=\boldsymbol{A}^{-1}\boldsymbol{y}$ 唯一确定. 而且一般也包括 $\boldsymbol{A}$ 不是正则矩阵的情况下,解存在的充分必要条件是 $\operatorname{rank}[\boldsymbol{A},\ \boldsymbol{y}]=\operatorname{rank}\boldsymbol{A}$(B-I-4 中定理 1),这时解是否也能用某个矩阵 $\boldsymbol{G}\in C^{n\times m}$ 表示成 $\boldsymbol{x}=\boldsymbol{Gy}$ 呢?

[定义 1] "对于 $\operatorname{rank}[\boldsymbol{A},\ \boldsymbol{y}]=\operatorname{rank}\boldsymbol{A}$ 中的任意 $\boldsymbol{y}\in C^{m}$,设 $\boldsymbol{x}$ 是(1)式的解,当有使 $\boldsymbol{x}=\boldsymbol{A}^{-}\boldsymbol{y}$ 成立的矩阵 $\boldsymbol{A}^{-}\in C^{n\times m}$* 存在时,$\boldsymbol{A}^{-}$ 称为 $\boldsymbol{A}$ 的广义逆矩阵."

在 $\operatorname{rank}[\boldsymbol{A},\ \boldsymbol{y}]=\operatorname{rank}\boldsymbol{A}$ 的条件下(1)式有解,但因其解不是唯一的,所以一般不能期望在定义 1 的意义上 $\boldsymbol{A}^{-}$ 也是唯一的. 因此至少必须先确定它的存在性.

(性质 1)

$\boldsymbol{A}^{-}$ 存在⇔有满足 $\boldsymbol{AA}^{-}\boldsymbol{A}=\boldsymbol{A}$ 的 $\boldsymbol{A}^{-}\in C^{n\times m}$ 存在

(证明) (⇒):满足 $\operatorname{rank}[\boldsymbol{A},\ \boldsymbol{y}]=\operatorname{rank}\boldsymbol{A}$ 的 $\boldsymbol{y}$,即(1)式有解时的 $\boldsymbol{y}$,可以表示成 $\boldsymbol{y}=\boldsymbol{Az}$,$\boldsymbol{z}\in C^{n}$,而且只能这样表示. 设 $\boldsymbol{Ax}=\boldsymbol{y}$ 的解可以表示成 $\boldsymbol{A}^{-}\boldsymbol{y}$,则 $\boldsymbol{AA}^{-}\boldsymbol{y}=\boldsymbol{y}$,即 $\boldsymbol{AA}^{-}\boldsymbol{Az}=\boldsymbol{Az}$. 因该式对于任意的 $\boldsymbol{z}\in C^{n}$ 均成立,则 $\boldsymbol{AA}^{-}\boldsymbol{A}=\boldsymbol{A}$. (⇐):

设 $\boldsymbol{x}\in C^{m}$ 为 $\boldsymbol{Ax}=\boldsymbol{y}$ 的解,则 $\boldsymbol{AA}^{-1}\boldsymbol{Ax}=\boldsymbol{Ax}$,即 $\boldsymbol{AA}^{-}\boldsymbol{y}=\boldsymbol{y}$ 成立,这意味着 $\tilde{\boldsymbol{x}}=\boldsymbol{A}^{-}\boldsymbol{y}$ 也是解.

证明完毕

(性质 2)

对于任意的 $\boldsymbol{A}\in C^{m\times n}$,至少有一个 $\boldsymbol{A}^{-}$ 存在,而且总是 $\operatorname{rank}\boldsymbol{A}^{-}\geqslant\operatorname{rank}\boldsymbol{A}$.

(证明) 若 $\boldsymbol{A}=\boldsymbol{0}_{m\times n}$,则 $^{\forall}\boldsymbol{X}\in C^{n\times m}$ 满足 $\boldsymbol{AXA}=\boldsymbol{A}$.

设 $\boldsymbol{A}\neq\boldsymbol{0}$,$\operatorname{rank}\boldsymbol{A}=r\,(r\geqslant 1)$. 将 $\boldsymbol{A}$ 表示成 $\boldsymbol{A}=\boldsymbol{BC}$,$\boldsymbol{B}\in C^{m\times r}$,$\boldsymbol{C}\in C^{r\times n}$,$\operatorname{rank}\boldsymbol{B}=\operatorname{rank}\boldsymbol{C}=r$(根据 B-I-3 中定理 4,总是可能的),令 $\boldsymbol{A}^{+}\triangleq\boldsymbol{C}^{*}(\boldsymbol{CC}^{*})^{-1}(\boldsymbol{B}^{*}\boldsymbol{B})^{-1}\boldsymbol{B}^{*}$. 因为这时 $\boldsymbol{AA}^{+}\boldsymbol{A}=\boldsymbol{A}$,则 $\boldsymbol{A}^{+}$ 是 $\boldsymbol{A}$ 的一种广义逆矩阵.

而且,因 $\operatorname{rank}\boldsymbol{A}=\operatorname{rank}\boldsymbol{AA}^{-}\boldsymbol{A}$,则 $\operatorname{rank}\boldsymbol{A}\leqslant\operatorname{rank}\boldsymbol{A}^{-}$.

证明完毕

由性质(1),(2)可见,对于任意 $\boldsymbol{A}$ 有在定义 1 上的广义逆矩阵 $\boldsymbol{A}^{-}$ 存在,作为和定义 1 等价的定义,所谓 $\boldsymbol{A}$ 的广义逆矩阵也可以定义成使得 $\boldsymbol{A}=\boldsymbol{AA}^{-}\boldsymbol{A}$ 成立的矩阵 $\boldsymbol{A}^{-}$. 由此很容易得出下列性质.

(性质 3)

$$(\boldsymbol{A}^{T})^{-}=(\boldsymbol{A}^{-})^{T},\ (\boldsymbol{A}^{*})^{-}=(\boldsymbol{A}^{-})^{*} \tag{2}$$

* 原文误印成 $\boldsymbol{A}^{-}\in C^{xnm}$.——译者注

对于(2)式的等号要多加注意．例如，第1式意味着$(\boldsymbol{A}^T)$的广义逆矩阵不是唯一的，而$(\boldsymbol{A}^-)^T$是其中的一个．

（性质4）

$$(\boldsymbol{A}\boldsymbol{A}^-)^2=\boldsymbol{A}\boldsymbol{A}^-,\ (\boldsymbol{A}^-\boldsymbol{A})^2=\boldsymbol{A}^-\boldsymbol{A} \tag{3}$$

（性质5）

若$\boldsymbol{A}\in C^{n\times n}$, $\det \boldsymbol{A}\neq 0$, 则$\boldsymbol{A}^-=\boldsymbol{A}^{-1}$且仅限于$\boldsymbol{A}^{-1}$．

（问题1）　试证明性质3～5．

下面的几个性质不象上面几个那样能够自明．

（性质6）

$\boldsymbol{A}\boldsymbol{A}^-\boldsymbol{B}=\boldsymbol{B}\Leftrightarrow$有满足$\boldsymbol{B}=\boldsymbol{A}\boldsymbol{C}$的矩阵$\boldsymbol{C}$存在$\Leftrightarrow$有满足$\boldsymbol{B}=\boldsymbol{A}\boldsymbol{A}^*\boldsymbol{D}$的矩阵$\boldsymbol{D}$存在．

（证明）　在第1个等价性中，$(\Rightarrow)$令$\boldsymbol{C}=\boldsymbol{A}^-\boldsymbol{B}$即可证明．$(\Leftarrow)$可以由$\boldsymbol{A}\boldsymbol{A}^-\boldsymbol{B}=\boldsymbol{A}\boldsymbol{A}^-\boldsymbol{A}\boldsymbol{C}=\boldsymbol{A}\boldsymbol{C}=\boldsymbol{B}$证明．在第2个等价性中，$(\Rightarrow)$可以令$\boldsymbol{D}=(\boldsymbol{A}^+)^*\boldsymbol{C}$得出．其中$\boldsymbol{A}^+$是在性质2的证明中所用过的$\boldsymbol{A}$的广义逆矩阵之一，如后面所述，对于$\boldsymbol{A}$是唯一确定的，而且特别是有$(\boldsymbol{A}^+\boldsymbol{A})^*=\boldsymbol{A}^+\boldsymbol{A}$的性质，在推导过程中使用了这个性质．$(\Leftarrow)$令$\boldsymbol{A}^*\boldsymbol{D}=\boldsymbol{C}$即可．由此可见，三个条件的等价性成立．

证明完毕

（性质7）

$$(\boldsymbol{A}\boldsymbol{A}^*+\boldsymbol{B})(\boldsymbol{A}\boldsymbol{A}^*+\boldsymbol{B})^-\boldsymbol{A}\boldsymbol{C}=\boldsymbol{A}\boldsymbol{C} \tag{4}$$

对于任意矩阵$\boldsymbol{C}$和任意准正定埃尔米特矩阵$\boldsymbol{B}$均成立．

（证明）

$$\boldsymbol{B}\geqslant\boldsymbol{0}\Rightarrow\text{可以写成 }\boldsymbol{B}=\boldsymbol{D}\boldsymbol{D}^*$$

$$\Rightarrow\boldsymbol{A}\boldsymbol{A}^*+\boldsymbol{B}=[\boldsymbol{A}\boldsymbol{D}]\begin{bmatrix}\boldsymbol{A}^*\\ \boldsymbol{D}^*\end{bmatrix}.$$

而且由性质6

$$\boldsymbol{A}\boldsymbol{C}=[\boldsymbol{A}\boldsymbol{D}]\begin{bmatrix}\boldsymbol{C}\\ \boldsymbol{0}\end{bmatrix}\Rightarrow\text{有使 }\boldsymbol{A}\boldsymbol{C}=[\boldsymbol{A}\boldsymbol{D}]\begin{bmatrix}\boldsymbol{A}^*\\ \boldsymbol{D}^*\end{bmatrix}\boldsymbol{E}=(\boldsymbol{A}\boldsymbol{A}^*+\boldsymbol{B})\boldsymbol{E}\text{ 的 }\boldsymbol{E}\text{存在．因此}$$

$$\begin{aligned}(\boldsymbol{A}\boldsymbol{A}^*+\boldsymbol{B})(\boldsymbol{A}\boldsymbol{A}^*+\boldsymbol{B})^-\boldsymbol{A}\boldsymbol{C}&=(\boldsymbol{A}\boldsymbol{A}^*+\boldsymbol{B})(\boldsymbol{A}\boldsymbol{A}^*+\boldsymbol{B})^-(\boldsymbol{A}\boldsymbol{A}^*+\boldsymbol{B})\boldsymbol{E}\\&=(\boldsymbol{A}\boldsymbol{A}^*+\boldsymbol{B})\boldsymbol{E}=\boldsymbol{A}\boldsymbol{C}\end{aligned}$$

证明完毕

（注意）

作为(4)式的特殊情况，若令$\boldsymbol{B}=\boldsymbol{0}$, $\boldsymbol{C}=\boldsymbol{I}_n$, 则

$$(\boldsymbol{A}\boldsymbol{A}^*)(\boldsymbol{A}\boldsymbol{A}^*)^-\boldsymbol{A}=\boldsymbol{A} \tag{5a}$$

成立．同时，无论$\boldsymbol{A}\boldsymbol{A}^*$的广义逆矩阵定成什么样，$\boldsymbol{A}^*(\boldsymbol{A}\boldsymbol{A}^*)^-\boldsymbol{A}$都不变，而且是埃尔米特矩阵．因此，取(5a)式的转置，

$$\boldsymbol{A}^*(\boldsymbol{A}\boldsymbol{A}^*)^-(\boldsymbol{A}\boldsymbol{A}^*)=\boldsymbol{A}^* \tag{5b}$$

也成立．

（问题2）　试证明$\boldsymbol{A}^*(\boldsymbol{A}\boldsymbol{A}^*)^-\boldsymbol{A}$是埃尔米特矩阵，而且无论定成什么样的广义逆矩阵，$\boldsymbol{A}^*(\boldsymbol{A}\boldsymbol{A}^*)^-\boldsymbol{A}$都不变．（提示：$\boldsymbol{A}[\boldsymbol{A}^*(\boldsymbol{A}\boldsymbol{A}^*)_1^-\boldsymbol{A}-\boldsymbol{A}^*(\boldsymbol{A}\boldsymbol{A})_2^-\boldsymbol{A}]=\boldsymbol{0}$．而且，对于任意

$x\in C^m,\ AA^*x=0\Rightarrow A^*x=0$)

(性质 8a)

对于埃尔米特矩阵 $A\in C^{n\times n}$,

$$A\triangleq\begin{array}{c c} \begin{array}{cc} \scriptstyle n_1 & \scriptstyle n-n_1 \end{array} & \\ \begin{bmatrix} \overbrace{A_{11}} & \overbrace{A_{12}} \\ A_{12}^* & A_{22} \end{bmatrix} & \begin{array}{l} \}n_1 \\ \}n-n_1 \end{array} \end{array} \geqslant 0\Leftrightarrow\begin{bmatrix} A_{22}\geqslant 0,\ A_{12}A_{22}^-A_{22}=A_{12} \\ A_{11}-A_{12}A_{22}^-A_{12}^*\geqslant 0 \end{bmatrix}$$

(证明) ($\Rightarrow$): 因 $A\geqslant 0$, 则由适当的 $B\in C^{(n-n_1)\times n_1}$, $C\in C^{(n-n_1)\times(n-n_1)}$ 可以表示成

$$A=\begin{bmatrix} B^* \\ C^* \end{bmatrix}[BC],$$

即

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{12}^* & A_{22} \end{bmatrix}=\begin{bmatrix} B^*B & B^*C \\ C^*B & C^*C \end{bmatrix}.$$

因此

$$A_{22}=C^*C\geqslant 0$$

$$A_{12}A_{22}^-A_{22}=B^*C(C^*C)^-C^*C=B^*C=A_{12},$$

而且

$$\begin{aligned} A_{11}-A_{12}A_{22}^-A_{12}^* &= A_{11}-A_{12}A_{22}^-A_{12}^*-(A_{12}A_{22}^-A_{12}^*-A_{12}A_{22}^-A_{12}^*) \\ &= B^*B-(B^*C)A_{22}^-A_{12}^*-A_{12}A_{22}^-(C^*B)+A_{12}A_{22}^-(C^*C)A_{22}^-A_{12}^* \\ &= (B^*-A_{12}A_{22}^-C^*)(B-CA_{22}^-A_{12}^*) \\ &= (B-CA_{22}^-A_{12}^*)^*(B-CA_{22}^-A_{12}^*)\geqslant 0 \end{aligned}$$

($\Leftarrow$):

$$\begin{bmatrix} I & -A_{12}A_{22}^- \\ 0 & I \end{bmatrix}A\begin{bmatrix} I & 0 \\ -A_{22}^-A_{12}^* & I \end{bmatrix}=\begin{bmatrix} A_{11}-A_{12}A_{22}^-A_{12}^* & 0 \\ 0 & A_{22} \end{bmatrix}\geqslant 0$$

$$\Rightarrow A\geqslant 0.$$

证明完毕

(性质 8b)

设 $A=A^*\in C^{n\times n}$, 则

$$\left.\begin{array}{l} A\triangleq\begin{bmatrix} A_{11} & A_{12} \\ A_{12}^* & A_{22} \end{bmatrix}\geqslant 0 \\ \operatorname{rank}A=\operatorname{rank}A_{22} \end{array}\right\}\Leftrightarrow\left\{\begin{array}{l} A_{22}\geqslant 0 \\ A_{12}A_{22}^-A_{22}=A_{12} \\ A_{11}-A_{12}A_{22}^-A_{12}^*=0 \end{array}\right.$$

(证明) ($\Rightarrow$): 设

$$\tilde{A}\triangleq\begin{bmatrix} I & -A_{12}A_{22}^- \\ 0 & I \end{bmatrix}A\begin{bmatrix} I & 0 \\ -A_{22}^-A_{12}^* & I \end{bmatrix}=\begin{bmatrix} A_{11}-A_{12}A_{22}^-A_{12}^* & 0 \\ 0 & A_{22} \end{bmatrix},$$

则 $\operatorname{rank}A=\operatorname{rank}\tilde{A}$. 因 $\operatorname{rank}A=\operatorname{rank}A_{22}$, 则必须 $A_{11}-A_{12}A_{22}^-A_{12}^*=0$. 在(性质 8a)中已证明, $A_{22}\geqslant 0$, $A_{12}A_{22}^-A_{22}=A_{12}$. ($\Leftarrow$)其逆亦可完全同样证明.

证明完毕

(性质 9a) (A^- 的一般形)

设 $U\in C^{n\times m}$ 为任意矩阵, A^- 是 A 的一个广义逆矩阵, 则

$$X=A^-+U-A^-AUAA^- \tag{6}$$

是 A 的广义逆矩阵, 而且 A 的任何广义逆矩阵都可以表示成这种形式.

（证明） 在(6)式两边各左、右乘以 $\boldsymbol{A}$, 因

$$\boldsymbol{AXA}=\boldsymbol{AA}^-\boldsymbol{A}+\boldsymbol{AUA}-\boldsymbol{AA}^-\boldsymbol{AUAA}^-\boldsymbol{A}=\boldsymbol{A}+\boldsymbol{AUA}-\boldsymbol{AUA}=\boldsymbol{A},$$

由性质 1, $\boldsymbol{X}$ 是 $\boldsymbol{A}$ 的广义逆矩阵. 设 $\boldsymbol{X}$ 为 $\boldsymbol{A}$ 的任意广义逆矩阵时,

$$\mathbf{0}=\boldsymbol{A}(\boldsymbol{X}-\boldsymbol{A}^-)\boldsymbol{A}\Rightarrow\boldsymbol{X}=\boldsymbol{A}^-+(\boldsymbol{X}-\boldsymbol{A}^-)-\boldsymbol{A}^-\boldsymbol{A}(\boldsymbol{X}-\boldsymbol{A}^-)\boldsymbol{AA}^-$$

成立. 即, 令 $\boldsymbol{U}=\boldsymbol{X}-\boldsymbol{A}^-$, (6)式成立.

证明完毕

（性质 9b）

设 $\boldsymbol{A}^-$ 为 $\boldsymbol{A}$ 的一个广义逆矩阵, 则对于任意的 $\boldsymbol{V}$, $\boldsymbol{W}\in C^{n\times m}$,

$$\boldsymbol{X}=\boldsymbol{A}^-+\boldsymbol{V}(\boldsymbol{I}_m-\boldsymbol{AA}^-)+(\boldsymbol{I}_n-\boldsymbol{A}^-\boldsymbol{A})\boldsymbol{W} \tag{7}$$

是 $\boldsymbol{A}$ 的广义逆矩阵, 而且 $\boldsymbol{A}$ 的任意广义逆矩阵都可以用该形式表示.

（证明） 由(7)式容易证明 $\boldsymbol{AXA}=\boldsymbol{A}$. 而且, 当 $\boldsymbol{X}$ 是 $\boldsymbol{A}$ 的任意广义逆矩阵时, 令 $\boldsymbol{V}=\boldsymbol{X}-\boldsymbol{A}^-$, $\boldsymbol{W}=\boldsymbol{XAA}^-$, 则对于 $\boldsymbol{V}$, $\boldsymbol{W}$(7)式成立.

证明完毕

（问题 3） 令 $\boldsymbol{U}=\boldsymbol{V}(\boldsymbol{I}-\boldsymbol{AA}^-)+(\boldsymbol{I}-\boldsymbol{A}^-\boldsymbol{A})\boldsymbol{W}$, 试证明(7)式可以写成(6)式的形式.

自反广义逆矩阵

上面讲过的广义逆矩阵 $\boldsymbol{A}^-$ 和普通逆矩阵 $\boldsymbol{A}^{-1}$ 有若干共同的性质, 同时也有许多必须注意的区别. 这是因为, 仅在 $\boldsymbol{A}$ 是方阵, $\det\boldsymbol{A}\neq0$ 时 $\boldsymbol{A}^{-1}$ 才有意义, 而 $\boldsymbol{A}^-$ 始终存在.

关于 $\boldsymbol{A}^{-1}$, 当然 $(\boldsymbol{A}^{-1})^{-1}=\boldsymbol{A}$ 成立, 与此相对应, 在 $\boldsymbol{AA}^-\boldsymbol{A}=\boldsymbol{A}$, $\boldsymbol{A}^-\boldsymbol{AA}^-=\boldsymbol{A}^-$ 成立(即 $\boldsymbol{A}$ 是 $(\boldsymbol{A}^-)^-$ 的一种)的条件下考虑 $\boldsymbol{A}$ 的广义逆矩阵.

[定义 2] "对于 $\boldsymbol{A}\in C^{m\times n}$, 使

$$\boldsymbol{AA}_r^-\boldsymbol{A}=\boldsymbol{A},\ \boldsymbol{A}_r^-\boldsymbol{AA}_r^-=\boldsymbol{A}_r^- \tag{8}$$

成立的矩阵 $\boldsymbol{A}_r^-\in C^{n\times m}$ 称为 $\boldsymbol{A}$ 的自反广义逆矩阵(reflexive generalized inverse)."

（性质 r-1）

设 $\boldsymbol{A}^-$ 为 $\boldsymbol{A}$ 的某个广义逆矩阵. 这时

$$\boldsymbol{A}^-\text{是自反的}\Leftrightarrow\operatorname{rank}(\boldsymbol{A}^-)=\operatorname{rank}\boldsymbol{A}.$$

（证明） ($\Rightarrow$): 由(8)式, $\operatorname{rank}\boldsymbol{A}^-\leqslant\operatorname{rank}\boldsymbol{A}$, $\operatorname{rank}\boldsymbol{A}\leqslant\operatorname{rank}\boldsymbol{A}^-$, 即 $\operatorname{rank}\boldsymbol{A}^-=\operatorname{rank}\boldsymbol{A}$. ($\Leftarrow$): 设 $\operatorname{rank}\boldsymbol{A}=\operatorname{rank}\boldsymbol{A}^-$, 由(8)式可得出 $\operatorname{rank}(\boldsymbol{AA}^-)=\operatorname{rank}\boldsymbol{A}^-$. 因此, 对于某个 $\boldsymbol{x}\in C^m$, $(\boldsymbol{AA}^-)\boldsymbol{x}=0$ 意味着 $\boldsymbol{A}^-\boldsymbol{x}=0$. 即

$$(\boldsymbol{AA}^-(\boldsymbol{I}_m-\boldsymbol{AA}^-))=\mathbf{0}\Rightarrow\boldsymbol{A}^-(\boldsymbol{I}_m-\boldsymbol{AA}^-)=\mathbf{0}.$$

证明完毕

（问题 4） 试证明 $\mathbf{0}_{m\times n}$ 的自反广义逆矩阵仅为 $\mathbf{0}_{m\times n}$.

（性质 r-2）

假定 $\boldsymbol{A}\neq\mathbf{0}$, 当 $\operatorname{rank}A=r$ 时, 将 $\boldsymbol{A}$ 分解成 $\boldsymbol{A}=\boldsymbol{BC}$, 其中假定 $\boldsymbol{B}\in C^{m\times r}$, $\boldsymbol{C}\in C^{r\times n}$, $\operatorname{rank}\boldsymbol{B}=\operatorname{rank}\boldsymbol{C}=r$. 当 $\boldsymbol{B}$ 的左逆矩阵, $\boldsymbol{C}$ 的右逆矩阵分别表示成 $\boldsymbol{B}_L^{-1}$, $\boldsymbol{C}_R^{-1}$($\boldsymbol{B}_L^{-1}\boldsymbol{B}=\boldsymbol{I}_r$, $\boldsymbol{CC}_R^{-1}=\boldsymbol{I}_r$)时

$$\boldsymbol{A}_r^-=\boldsymbol{C}_R^{-1}\boldsymbol{B}_L^{-1} \tag{9}$$

是 $\boldsymbol{A}$ 的自反广义逆矩阵, 而且任何自反广义逆矩阵都可以用该形式表示.

(注意)

因 $\boldsymbol{B}, \boldsymbol{C}$ 有最大秩，则 $\boldsymbol{B}_L^{-1}$, $\boldsymbol{C}_R^{-1}$ 必定存在. 例如, $(\boldsymbol{B}^*\boldsymbol{B})^{-1}\boldsymbol{B}^*$, $\boldsymbol{C}^*(\boldsymbol{C}\boldsymbol{C}^*)^{-1}$ 分别为 $\boldsymbol{B}_L^{-1}$, $\boldsymbol{C}_R^{-1}$ 中的一个. 因 $\boldsymbol{B}_L^{-1}$, $\boldsymbol{C}_R^{-1}$ 存在,则对于任意 $\boldsymbol{A}$, $\boldsymbol{A}_r^{-1}$ 也存在(但不是唯一的).

(证明) 由(9)式得, $\boldsymbol{A}\boldsymbol{A}_r^-\boldsymbol{A}=\boldsymbol{B}\boldsymbol{C}\boldsymbol{C}_R^{-1}\boldsymbol{B}_L^{-1}\boldsymbol{B}\boldsymbol{C}=\boldsymbol{B}\boldsymbol{C}=\boldsymbol{A}$, 同样,因 $\boldsymbol{A}_r^{-1}\boldsymbol{A}\boldsymbol{A}_r^{-1}=\boldsymbol{A}_r^-$ 成立,则(9)式的 $\boldsymbol{A}_r^-$ 是自反广义逆矩阵.

设 $\boldsymbol{A}_r^-$ 为 $\boldsymbol{A}$ 的任意自反广义逆矩阵,则 $\boldsymbol{A}\boldsymbol{A}_r^-\boldsymbol{A}=\boldsymbol{A}$, 即 $\boldsymbol{B}\boldsymbol{C}\boldsymbol{A}_r^-\boldsymbol{B}\boldsymbol{C}=\boldsymbol{B}\boldsymbol{C}$. 将该式两边分别左乘以 $\boldsymbol{B}_L^{-1}$,右乘以 $\boldsymbol{C}_R^{-1}$,得 $\boldsymbol{C}\boldsymbol{A}_r^-\boldsymbol{B}=\boldsymbol{I}_r$. 该式意味着 $\boldsymbol{A}_r^-\boldsymbol{B}$ 是 $\boldsymbol{C}$ 的右逆矩阵, $\boldsymbol{C}\boldsymbol{A}_r^-$ 是 $\boldsymbol{B}$ 的左逆矩阵. 因此,写成 $\boldsymbol{A}_r^-\boldsymbol{B}=\widetilde{\boldsymbol{C}}_R^{-1}$, $\boldsymbol{C}\boldsymbol{A}_r^-=\widetilde{\boldsymbol{B}}_L^{-1}$. 代入 $\boldsymbol{A}_r^-=\boldsymbol{A}_r^-\boldsymbol{A}\boldsymbol{A}_r^-$ 后, 得 $\boldsymbol{A}_r^-=\boldsymbol{A}_r^-\boldsymbol{B}\boldsymbol{C}\boldsymbol{A}_r=\widetilde{\boldsymbol{C}}_R^{-1}\widetilde{\boldsymbol{B}}_L^{-1}$.

证明完毕

伪逆矩阵

如(性质r-2)中所述,当 $\boldsymbol{A}$ 用有最大秩阵之积 $\boldsymbol{A}=\boldsymbol{B}\boldsymbol{C}$ 表示时,容易证明

$$\boldsymbol{A}^+\triangleq\boldsymbol{C}^*(\boldsymbol{C}\boldsymbol{C}^*)^{-1}(\boldsymbol{B}^*\boldsymbol{B})^{-1}\boldsymbol{B}^* \tag{10}$$

是 $\boldsymbol{A}$ 的自反广义逆矩阵. 但是应该注意, 虽然 $\boldsymbol{B}, \boldsymbol{C}$ 不是由 $\boldsymbol{A}$ 唯一确定, 但是(10)式给出唯一的矩阵 $\boldsymbol{A}^+$ 1). 即, 对于 $\boldsymbol{A}$, $\boldsymbol{A}^+$ 有唯一确定的特征. 其唯一性在B-I-3[系]中已证明. 这个 $\boldsymbol{A}^+$ 称为 $\boldsymbol{A}$ 的伪逆矩阵(pseudo-inverse)或穆尔-彭罗斯逆矩阵(Moore-Penrose inverse).

利用 $\boldsymbol{A}^+$ 的唯一性及(10)式,容易得出下列性质.

(性质 p-1)

$$(\boldsymbol{A}^+)^+=\boldsymbol{A}.$$

(性质 p-2)

$$(\boldsymbol{A}^*)^+=(\boldsymbol{A}^+)^*,\ (\boldsymbol{A}\boldsymbol{A}^*)^+=(\boldsymbol{A}^+)^*\boldsymbol{A}^+.$$

(性质 p-3)

$$(\boldsymbol{A}\boldsymbol{A}^+)^*=\boldsymbol{A}\boldsymbol{A}^+,\ (\boldsymbol{A}^+\boldsymbol{A})^*=\boldsymbol{A}^+\boldsymbol{A}.$$

(证明) 由(10)式,因

$$\boldsymbol{A}\boldsymbol{A}^+=\boldsymbol{B}(\boldsymbol{B}^*\boldsymbol{B})^{-1}\boldsymbol{B}^*,\ \boldsymbol{A}^+\boldsymbol{A}=\boldsymbol{C}^*(\boldsymbol{C}\boldsymbol{C}^*)^{-1}\boldsymbol{C},$$

则

$$(\boldsymbol{A}\boldsymbol{A}^+)^*=\boldsymbol{A}\boldsymbol{A}^+,(\boldsymbol{A}^+\boldsymbol{A})^*=\boldsymbol{A}^+\boldsymbol{A}. \tag{11}$$

证明完毕

(注意)

我们知道, $\boldsymbol{A}^+$ 除了是自反广义逆矩阵外还满足(性质 p-3), 即 $\boldsymbol{A}^+$ 由下列四个条件

$$\boldsymbol{A}\boldsymbol{A}^+\boldsymbol{A}=\boldsymbol{A},\ \boldsymbol{A}^+\boldsymbol{A}\boldsymbol{A}^+=\boldsymbol{A}^+,\ (\boldsymbol{A}\boldsymbol{A}^+)^*=\boldsymbol{A}\boldsymbol{A}^+,\ (\boldsymbol{A}^+\boldsymbol{A})^*=\boldsymbol{A}^+\boldsymbol{A} \tag{12}$$

唯一确定,这四个条件也可以作为 $\boldsymbol{A}^+$ 的定义[129].

(性质 p-4)

$$\boldsymbol{A}^+=(\boldsymbol{A}^*\boldsymbol{A})^+\boldsymbol{A}^*=\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{A}^+.$$

(证明) 设 $\boldsymbol{A}=\boldsymbol{B}\boldsymbol{C}$, 则 $\boldsymbol{A}^*\boldsymbol{A}$ 的最大秩分解可写成 $\boldsymbol{A}^*\boldsymbol{A}=\boldsymbol{C}^*(\boldsymbol{B}^*\boldsymbol{B}\boldsymbol{C})$, 则

$$\begin{aligned}(\boldsymbol{A}^*\boldsymbol{A})^+&=(\boldsymbol{B}^*\boldsymbol{B}\boldsymbol{C})^*(\boldsymbol{B}^*\boldsymbol{B}\boldsymbol{C}\boldsymbol{C}^*\boldsymbol{B}^*\boldsymbol{B})^{-1}(\boldsymbol{C}\boldsymbol{C}^*)^{-1}\boldsymbol{C}\\&=\boldsymbol{C}^*(\boldsymbol{C}\boldsymbol{C}^*)^{-1}(\boldsymbol{B}^*\boldsymbol{B})^{-1}(\boldsymbol{C}\boldsymbol{C}^*)^{-1}\boldsymbol{C}.\end{aligned}$$

1) 对于 $\boldsymbol{A}=\boldsymbol{0}_{m\times n}$, 假定 $\boldsymbol{A}^+\triangleq\boldsymbol{0}_{n\times m}$.

因此

$$(A^*A)^+A^*=C^*(CC^*)^{-1}(B^*B)^{-1}(CC^*)^{-1}C(C^*B^*)=A^+.$$

$$A^*(AA^*)^+=A^+$$

也可以同样证明.

证明完毕

(问题 5) 试证明性质 p-1, p-2.

(问题 6) 虽然 A 的最大秩分解 $A=BC$ 不是唯一的，但是(11)式右边的

$$B(B^*B)^{-1}B^* \quad 及 \quad C^*(CC^*)^{-1}C$$

是唯一的，试证明之.

(问题 7) 试证明下列性质.

(i) $(A^*A)^+=A^+(A^*)^+$, $(AA^*)^+=(A^*)^+A^+$

(ii) $(A^*A)^+=A^+(AA^*)^+A=A^*(AA^*)^+(A^*)^+$

(iii) $A^*=A \Rightarrow (A^2)^+=(A^+)^2$, $A^2(A^2)^+=(A^2)^+A^2=AA^+$, $A^+A^2=A^2A^+$

(iv) $A^*=A \Rightarrow AA^+=A^+A$

(v) $AA^+=(AA^*)(AA^*)^+=(AA^*)^+(AA^*)$

$A^+A=(A^*A)(A^*A)^+=(A^*A)^+(A^*A)$

(性质 p-5)

对于 $A\in C^{m\times n}$, $B\in C^{n\times p}$, 一般 $(AB)^+=B^+A^+$ 不成立. 设 $B_1\triangleq A^+AB$, $A_1\triangleq AB_1B_1^+$ 时，一般

$$(AB)^+=B_1^+A_1^+$$

(证明) 因 $A_1B_1=AB_1B_1^+B_1=AB_1=AA^+AB=AB$, 则只要证明 $B_1^+A_1^+$ 是 A_1B_1 的伪逆矩阵即可. 设 $X\triangleq B_1^+A_1^+$, $Y\triangleq A_1B_1$, 利用上面所示的关系 $A_1B_1=AB_1=AB$, 则容易证明 $XYX=X$, $YXY=Y$. 若能证明 YX, XY 是埃尔米特矩阵，由(12)式则可以证明 $X=Y^+$.

在 $$YX=A_1(B_1B_1^+A_1^+)=AB_1B_1^+(B_1B_1^+A_1^+)=AB_1B_1^+A_1^+=A_1A_1^+$$

中，因 $A_1A_1^+$ 是埃尔米特矩阵，则 YX 也是埃尔米特矩阵.

$$XY=B_1^+(A_1+A_1)B_1=B_1^*(A_1+A_1)^*B_1=B_1^+(A_1^+A_1B_1B_1^+)^*B_1$$
$$=B_1^+(B_1B_1^+A_1^+A_1)B_1=B_1^+(A^+ABB_1^+A_1^+A_1)B_1,$$

同样, $XY=B^+B_1B_1^+B_1$. 这就证明 XY 是埃尔米特矩阵.

证明完毕

(注意)

若 $\operatorname{rank} A=\operatorname{rank} B=n$, 因 $C\triangleq AB$ 是 C 的最大秩分解，则

$$C^+\triangleq(AB)^+=B^*(BB^*)^{-1}(A^*A)^{-1}A^*=B^+A^+$$

成立. 我们知道[129~130], 一般 $(AB)^+=B^+A^+$ 的条件不自明, $A^+AB(AB)^*=B(AB)^*$ 及 $BB^+A^*AB=A^*AB$ 同时成立是充分必要条件.

(问题 8) 试证明 $(A\otimes B)^+=(A^+\otimes B^+)$.

(性质 p-6)

设矩阵序列 $A_k\in C^{m\times n}$ 具有极限 $A\triangleq\lim\limits_{k\to\infty}A_k$. 若极限 $X\triangleq\lim\limits_{k\to\infty}A_k^+$ 存在，则 $X=A^+$

(注意)

A^+ 必定存在，但是，$\lim\limits_{k\to\infty} A_k^+$ 不一定给出 A^+. 例如，设 $A_k \triangleq \left[\frac{1}{k}\right]$，则

$$A = \lim_{k\to\infty} A_k = \lim_{k\to\infty}\left[\frac{1}{k}\right] = [0].$$

由定义，$A^+ = [0]$，另一方面，$A_k^+ = [k]$，$A_k^+ \to \infty (k\to\infty)$.

当 A 是正则矩阵时，若将 A 划分成分块矩阵，则 A^{-1} 的计算很方便(B-I-1 定理 3). 在伪逆矩阵的情况下，也可以得出与其相当的结果. 因相当复杂故这里省略，可参考有关文献[129].

A^+ 的各种表示

当给出 $A \in C^{m\times n}$ 时，A^+ 可由(10)式求出，但是除此之外还有各种各样的表示方法. 当 A 是对角形，即 $A = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ 时，容易证明 $A^+ = \mathrm{diag}(\beta_1, \beta_2, \cdots, \beta_n)$，其中假定 $\beta_i = \lambda_i^{-1} (\lambda_i \neq 0)$，$\beta_i = 0 (\lambda_i = 0)$. 现在将埃尔米特矩阵 A^*A 表示成(参照 B-I-11 定理 10)

$$A^*A = U\Lambda U^*, \tag{13}$$

式中 U 是酉阵，它的列由 A^*A 的特征向量构成. 而且 Λ 是以 A^*A 的特征值为主对角线上元素的实对角矩阵. 因 A^*A 是准正定的，则其特征值都是非负的(参照 B-I-11 中系 19, 20). 设 Λ^+ 如上述那样确定，则

$$(A^*A)^+ = U\Lambda^+U^* \tag{14}$$

成立(试证明之). 此外，由(性质 p-4)得 $A^+ = (A^*A)^+A^*$，将(14)式代入便可求出 A^+. 把它再稍微详细计算一下，可以得到和一般矩阵函数 $f(A)$ 的拉格朗日-西勒维斯特内插多项式(参照 B-I-13 中(25)式)相类似的 A^+ 的表示式. 即表示成

$$\Lambda = \begin{bmatrix} \lambda_1^2 & & & & & \\ & \lambda_2^2 & & & & \\ & & \ddots & & & \\ & & & \lambda_r^2 & & \\ & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{bmatrix} \tag{15}$$

时，(13)，(14)式可以写成谱分解形(B-I-11 中定理 5)

$$A^*A = \sum_{j=1}^{r} \lambda_j^2 E_j,\quad (A^*A)^+ = \sum_{j=1}^{r} \lambda_j^{-2} E_j. \tag{16}$$

式中 $E_j^2 = E_j (j = 1, 2, \cdots, r)$，$E_jE_k = 0 (j \neq k)$. 现引入多项式

$$p_i(\lambda) \triangleq (\lambda - \lambda_1^2)\cdots(\lambda - \lambda_{i-1}^2)(\lambda - \lambda_{i+1}^2)\cdots(\lambda - \lambda_r^2),$$

则

$$p_i(A^*A) = \sum_{j=1}^{r} p_i(\lambda_j^2) E_j, \tag{17}$$

因 $p_i(\lambda_j^2) = 0$，$\forall j \neq i$，故(17)式变成

$$p_i(A^*A) = p_i(\lambda_i^2) E_i.$$

由该式解出 E_i 后代入(16)式，便可得到 A^+ 的如下表示式

$$A^+ = \sum_{j=1}^{r} \lambda_j^{-2} E_j A^* = \sum_{j=1}^{r} \lambda_j^{-2} \frac{p_j(A^*A)}{p_j(\lambda_j^2)} A^* = \sum_{j=1}^{r} \lambda_j^{-2} \frac{\prod\limits_{i\neq j}(A^*A - \lambda_i^2 I)}{\prod\limits_{i\neq j}(\lambda_j^2 - \lambda_i^2)} A^* \tag{18}$$

利用 A 的所谓奇异值分解(singular decomposition),也可以得到 A^+ 的一个标准形.现在在(13),(15)式中令

$$U \triangleq [\overbrace{P}^{r} | \overbrace{\tilde{P}}^{n-r}]\}n, \tag{19}$$

$$L \triangleq \begin{bmatrix} \lambda_1^2 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_r^2 \end{bmatrix} > \mathbf{0}, \tag{20}$$

则可写成

$$A^*A = U\Lambda U^* = PLP^*. \tag{21}$$

注意,因 U 是酉阵,则 $P^*P = I_r$. 设 L 的正定平方根矩阵为 $L^{1/2}$,则显然 $PL^{1/2}P^*$ 是 A^*A 的准正定平方根矩阵,并且容易证明,和(14)式一样,其伪逆矩阵为

$$[(A^*A)^{1/2}]^+ = PL^{-1/2}P^*. \tag{22}$$

此外,利用(问题7)(iii),(iv),(v)的结果,得

$$A^+A = (A^*A)^+(A^*A) = [(A^*A)^{1/2}]^+(A^*A)^{1/2}. \tag{23}$$

将(22),(23)式代入 $A = AA^+A$,则

$$A = A[(A^*A)^{1/2}]^+(A^*A)^{1/2} = A[(A^*A)^{1/2}]^+PL^{1/2}P^* = Q^*L^{1/2}P^*, \tag{24}$$

式中

$$Q^* \triangleq A[(A^*A)^{1/2}]^+P. \tag{25}$$

(24)式称为 A 的奇异值分解形.

由(25)式,得

$$\begin{aligned} Q^*LQ &= A[(A^*A)^{1/2}]^+PLP^*[(A^*A)^{1/2}]^+A^* \\ &= A[(A^*A)^{1/2}]^+A^*A[(A^*A)^{1/2}]^+A^* \\ &= A(A^+AA^+A)A^* = AA^* \end{aligned} \tag{26}$$

利用(26)及(22)式,可见关于 Q 有下列性质

$$\begin{aligned} QQ^* &= P^*[(A^*A)^{1/2}]^+A^*A[(A^*A)^{1/2}]^+P \\ &= P^*[PL^{-1/2}P^*]A^*A[PL^{-1/2}P^*]P = I_r. \end{aligned} \tag{27}$$

再者,因奇异值分解也是 A 的最大秩分解,则由(24)式立即可求出 A^+. 即,设 $C \triangleq L^{1/2}P^*$, $B \triangleq Q^*$,利用(10)式 A^+ 可表示成

$$A^+ = (PL^{1/2})(L^{1/2}P^*PL^{1/2})^{-1}(QQ^*)^{-1}Q = PL^{-1/2}Q \tag{28}$$

便得到 A^+ 的一个标准形.

将以上结果归纳,得如下定理.

[定理 1] 当 $\operatorname{rank} A = r$ 时,设 A^*A 的正特征值(同样,或 AA^* 的正特征值)为 λ_1^2, λ_2^2, $\cdots$, λ_r^2. 这时,A^+ 可用(18)式表示. 而且,当 $L \triangleq \operatorname{diag}(\lambda_1^2, \lambda_2^2, \cdots, \lambda_r^2) > \mathbf{0}$,设 $L^{1/2}$ 为 L 的正定平方根矩阵,即 $L^{1/2} \triangleq \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_r) > \mathbf{0}$ 时,也可以表示成 $A^+ = PL^{-1/2}Q$. 其中,$P \in C^{n\times r}$, $Q \in C^{r\times m}$ 是满足下列关系的矩阵,

$$A^*A = PLP^*,\ P^*P = I_r,$$

$$AA^* = Q^*LQ,\ QQ^* = I_r.$$

(问题 9) 试证明 $AA^+ = Q^*Q$, $A^+A = PP^*$.

作为 A^+ 的第 4 种表示式,有下列经常用的形式.

[定理 2]

$$A^+=\lim_{\delta\to 0}(A^*A+\delta^2 I_n)^{-1}A^*=\lim_{\delta\to 0}A^*(AA^*+\delta^2 I_m)^{-1}. \tag{29}$$

(问题 10) 当 U 是埃尔米特矩阵, Λ 是对角矩阵时, 试利用(20)式证明 $(U\Lambda U^*)^+=U\Lambda^+U^*$ 成立. 当 U 只是正则矩阵时, 一般 $(U\Lambda U^*)^+=(U^*)^{-1}\Lambda^+U^{-1}$ 不成立, 试举出不成立的例子加以证明.

因定理 2 的证明要用到下列引理, 故先来推出该引理.

[引理][130] 对于任意埃尔米特矩阵 $A\in C^{n\times n}$,

$$AA^+(=A^+A)=\lim_{\delta\to 0}(A+\delta I)^{-1}A=\lim_{\delta\to 0}A(A+\delta I)^{-1}.$$

(引理的证明) 对于足够小的 $|\delta|>0$, 因 $(A+\delta I)$ 是正则的, 则 $(A+\delta I)^{-1}$ 存在.

因 A 是埃尔米特矩阵, 则可由酉阵 U 表示成

$$A=U\Lambda U^*,\ \Lambda=\mathrm{diag}[\lambda_1(A),\ \cdots,\ \lambda_n(A)].$$

注意, $A=AA^+A=A^2A^+$, 可见

$$(A+\delta I)^{-1}A=(A+\delta I)^{-1}A^2A^+=U(\Lambda+\delta I)^{-1}\Lambda^2U^*A^+$$

成立. 因

$$\lim_{\delta\to 0}(\Lambda+\delta I)^{-1}\Lambda^2=\lim_{\delta\to 0}\mathrm{diag}[(\lambda_1+\delta)^{-1}\lambda_1^2,\ \cdots,\ (\lambda_n+\delta)^{-1}\lambda_n^2]=\mathrm{diag}(\lambda_1,\ \cdots,\ \lambda_n)=\Lambda$$

则

$$\lim_{\delta\to 0}(A+\delta I)^{-1}A=U\Lambda U^*A^+=AA^+.$$

而 $\lim\limits_{\delta\to 0}A(A+\delta I)=AA^+$ 也可以同样推导出.

证明完毕

(定理 2 的证明)

注意到 $A^*=A^*(A^*)^+A^*=A^*(AA^+)^*=A^*AA^+$, 则

$$\lim_{\delta\to 0}(A^*A+\delta^2 I_n)^{-1}A^*=\lim_{\delta\to 0}(A^*A+\delta^2 I_n)^{-1}A^*AA^+.$$

将引理用于上式, 得

$$=(A^*A)(A^*A)^+A^+$$

再利用问题 7 中(v)

$$=A^+AA^+=A^+,$$

即证明了(29)式中的第 1 式.

$A^+=\lim\limits_{\delta\to 0}A^*(AA^*+\delta^2 I_m)^{-1}$ 也可以同样证明.

证明完毕

在线性方程组中的应用

在线性方程组 $Ax=y$ 的解存在的情况下, 它的一个特解可以表示成 $x=A^-y$, 当然这不是一般解. 以下将要讨论一般解怎样用 A^- 表示, 而且在解不存在的情况下(即 $\mathrm{rank}[A|y]\neq\mathrm{rank}[A]$, 也可以用 A^+ 求得在某种意义上最好的近似解, 等等.

[定理 3] "关于 x 的线性方程组

$$Ax=y,\ A\in C^{m\times n},\ y\in C^m \tag{30}$$

下列结果成立.

[I] $\boldsymbol{Ax}=\boldsymbol{y}$ 有解的情况

(i) $\boldsymbol{Ax}=\boldsymbol{y}$ 有解$\Leftrightarrow \boldsymbol{AA^-y}=\boldsymbol{y}$

(ii) 设 $\boldsymbol{A}^-$ 为 $\boldsymbol{A}$ 的任意广义逆矩阵，将它固定时，解的一般形可以表示成

$$\boldsymbol{x}=\boldsymbol{A}^-\boldsymbol{y}+(\boldsymbol{I}_n-\boldsymbol{A}^-\boldsymbol{A})\boldsymbol{z} \tag{31}$$

式中 $\boldsymbol{z}\in C^n$ 是任意向量.

(iii) $\hat{\boldsymbol{x}}=\boldsymbol{A}^+\boldsymbol{y}$ 是 $\boldsymbol{Ax}=\boldsymbol{y}$ 的所有解中有最小范数(对于欧几里得范数)的解.

[II] $\boldsymbol{Ax}=\boldsymbol{y}$ 没有解的情况

(iv) $\hat{\boldsymbol{x}}=\boldsymbol{A}^+\boldsymbol{y}$ 是所有最小二乘近似解中有最小范数的向量. 即，在满足下列关系

$$\|\boldsymbol{Ax}-\boldsymbol{y}\|\leqslant\|\boldsymbol{Az}-\boldsymbol{y}\|,\ \forall\boldsymbol{z}\in C^n \tag{32}$$

的最小二乘近似解 $\boldsymbol{x}$ 的集合中，对于 $\hat{\boldsymbol{x}}=\boldsymbol{A}^+\boldsymbol{y}$,

$$\|\hat{\boldsymbol{x}}\|\leqslant\|\boldsymbol{x}\| \tag{33}$$

成立. 式中 $\|\cdot\|$ 是欧几里得范数.

(证明)

(i) ($\Rightarrow$): 当 $\boldsymbol{Ax}=\boldsymbol{y}$ 对于某个 $\boldsymbol{x}\in C^n$ 成立时,

$$\boldsymbol{AA^-y}=\boldsymbol{AA^-Ax}=\boldsymbol{Ax}=\boldsymbol{y}.$$

($\Leftarrow$): $\boldsymbol{AA^-y}=\boldsymbol{y}$* 表明 $\boldsymbol{x}\triangleq\boldsymbol{A^-y}$ 满足 $\boldsymbol{Ax}=\boldsymbol{y}$.

(ii) 在(31)式**两边左乘以 $\boldsymbol{A}$, 则

$$\boldsymbol{Ax}=\boldsymbol{AA^-y}-\boldsymbol{A}(\boldsymbol{I}_n-\boldsymbol{A^-A})\boldsymbol{z}=\boldsymbol{AA^-y}.$$

由(i)，因 $\boldsymbol{AA^-y}=\boldsymbol{y}$, 显然(31)式的 $\boldsymbol{x}$ 是(30)式的解. 而且，当设 $\tilde{\boldsymbol{x}}$ 为任意解时，若令 $\boldsymbol{z}=\tilde{\boldsymbol{x}}-\boldsymbol{A^-y}$, 因

$$\begin{aligned}(\boldsymbol{I}_n-\boldsymbol{A^-A})\boldsymbol{z}&=(\boldsymbol{I}_n-\boldsymbol{A^-A})(\tilde{\boldsymbol{x}}-\boldsymbol{A^-y})\\&=\tilde{\boldsymbol{x}}-\boldsymbol{A^-y}-\boldsymbol{A^-A}\tilde{\boldsymbol{x}}+\boldsymbol{A^-AA^-y}=\tilde{\boldsymbol{x}}-\boldsymbol{A^-y}-\boldsymbol{A^-y}+\boldsymbol{A^-y},\end{aligned}$$

则(31)式成立.

(iii), (iv) 用欧几里得范数时

$$\begin{aligned}\|\boldsymbol{Az}-\boldsymbol{y}\|^2&=\|(\boldsymbol{AA}^+-\boldsymbol{I}_m)\boldsymbol{y}+\boldsymbol{A}(\boldsymbol{z}-\boldsymbol{A}^+\boldsymbol{y})\|^2\\&=\|(\boldsymbol{AA}^+-\boldsymbol{I}_m)\boldsymbol{y}\|^2+\|\boldsymbol{A}(\boldsymbol{z}-\boldsymbol{A}^+\boldsymbol{y})\|^2\\&\quad+\boldsymbol{y}^*(\boldsymbol{AA}^+-\boldsymbol{I}_m)^*\boldsymbol{A}(\boldsymbol{z}-\boldsymbol{A}^+\boldsymbol{y})+(\boldsymbol{z}-\boldsymbol{A}^+\boldsymbol{y})^*\boldsymbol{A}^*(\boldsymbol{AA}^+-\boldsymbol{I}_m)\boldsymbol{y}.\end{aligned}$$

由性质 p-3, $(\boldsymbol{AA}^+)^*\boldsymbol{A}=\boldsymbol{A}$, $\boldsymbol{A}^*(\boldsymbol{AA}^+)=\boldsymbol{A}^*$ 成立，将其代入上式消去最后两项. 因此

$$\begin{aligned}\|\boldsymbol{Az}-\boldsymbol{y}\|^2&=\|(\boldsymbol{AA}^+-\boldsymbol{I}_m)\boldsymbol{y}\|^2+\|\boldsymbol{A}(\boldsymbol{z}-\boldsymbol{A}^+\boldsymbol{y})\|^2\\&\geqslant\|(\boldsymbol{AA}^+-\boldsymbol{I}_m)\boldsymbol{y}\|^2,\ \forall\boldsymbol{z}\in C^n,\ \forall\boldsymbol{y}\in C^m\end{aligned}$$

成立. 特别是当 $\boldsymbol{y}$ 固定时，若选择 $\boldsymbol{z}$ 使得 $\boldsymbol{Az}=\boldsymbol{AA}^+\boldsymbol{y}$, 则 $\|\boldsymbol{Az}-\boldsymbol{y}\|^2$ 与其最小值 $\|(\boldsymbol{AA}^+-\boldsymbol{I}_m)\boldsymbol{y}\|^2$ 一致. 此外，由(ii)，设 $\boldsymbol{w}\in C^n$ 为任意向量，则 $\boldsymbol{Az}=\boldsymbol{AA}^+\boldsymbol{y}$ 的一般解可以表示成

$$\boldsymbol{z}=\boldsymbol{A}^+(\boldsymbol{AA}^+\boldsymbol{y})+(\boldsymbol{I}_n-\boldsymbol{A}^+\boldsymbol{A})\boldsymbol{w}=\boldsymbol{A}^+\boldsymbol{y}+(\boldsymbol{I}_n-\boldsymbol{A}^+\boldsymbol{A})\boldsymbol{w}.$$

但是因

$$\|\boldsymbol{z}\|^2=\|\boldsymbol{A}^+\boldsymbol{y}\|^2+\|(\boldsymbol{I}_n-\boldsymbol{A}^+\boldsymbol{A})\boldsymbol{w}\|^2 \tag{34}$$

(试证明)，则表明在使 $\|\boldsymbol{Az}-\boldsymbol{y}\|^2$ 的值为最小的 $\boldsymbol{z}$ 值中，$\hat{\boldsymbol{z}}=\boldsymbol{A}^+\boldsymbol{y}$ 有最小范数. 当然，在

---

* 原文误印成 $\boldsymbol{AA^-y}$. ——译者注

** 原文误为在(2)式. ——译者注

$\boldsymbol{Ax}=\boldsymbol{y}$ 有解的情况下，由(i)，因 $\boldsymbol{AA}^+\boldsymbol{y}=\boldsymbol{y}$ 成立，则 $\hat{\boldsymbol{z}}=\boldsymbol{A}^+\boldsymbol{y}$ 是满足 $\boldsymbol{Az}=\boldsymbol{y}$ 的 $\boldsymbol{z}$ 中有最小范数的向量.

证明完毕

(注意)

(i) 在(31)式中，$\boldsymbol{A}^-\boldsymbol{y}$ 表示方程式 $\boldsymbol{Ax}=\boldsymbol{y}$ 的一个特解. 显然，$(\boldsymbol{I}_n-\boldsymbol{A}^-\boldsymbol{A})\boldsymbol{z}$, $\boldsymbol{z}\in C^n$ 是令 $\boldsymbol{y}=\boldsymbol{0}$ 时 $\boldsymbol{Ax}=\boldsymbol{0}$ 的解. 满足 $\boldsymbol{Ax}=\boldsymbol{0}$ 的 $\boldsymbol{x}\in C^n$ 的集合，是 C^n 的部分空间(参看B-O-2)，称为 $\boldsymbol{A}$ 的零空间(null space). 定理3还意味着，适当地设定 $\boldsymbol{z}$, $\boldsymbol{A}$ 的零空间 $N(\boldsymbol{A})$ 中的任意向量可以表示成 $(\boldsymbol{I}_n-\boldsymbol{A}^-\boldsymbol{A})\boldsymbol{z}$ 的形式.

(ii) 一般，用 $\boldsymbol{x}=\boldsymbol{Az}$, $\boldsymbol{z}\in C^n$(任意)形式表示的向量 $\boldsymbol{x}$ 的集合，换句话说，用 $\boldsymbol{A}$ 的列线性表出的向量 $\boldsymbol{x}$ 的集合，称为 $\boldsymbol{A}$ 的值域(range). 值域仍然是 C^m 的部分空间. 而且，当给出某个向量 $\boldsymbol{y}\in C^m$ 时，满足

$$\|\boldsymbol{y}-\tilde{\boldsymbol{x}}\|=\min_{\boldsymbol{x}\in R(\boldsymbol{A})}\|\boldsymbol{y}-\boldsymbol{x}\|=\min_{\boldsymbol{z}\in C^n}\|\boldsymbol{y}-\boldsymbol{Az}\|$$

的向量 $\tilde{\boldsymbol{x}}$ 称为 $\boldsymbol{y}$ 到 $\boldsymbol{A}$ 的值域 $R(\boldsymbol{A})$ 的正交投影(orthogonal projection). 这里假定 $\|\cdot\|$ 是欧几里得范数. 在 $R(\boldsymbol{A})$ 中最接近 $\boldsymbol{y}$ 的向量 $\tilde{\boldsymbol{x}}$ 是 $\boldsymbol{y}$ 的投影.

定理3 (iv)表明，

$$\tilde{\boldsymbol{x}}=\boldsymbol{AA}^+\boldsymbol{y}$$

给出 $\boldsymbol{y}$ 到 $R(\boldsymbol{A})$ 的正交投影. $\tilde{\boldsymbol{x}}$ 本身由 $\boldsymbol{y}$ 唯一确定，而一般

$$\boldsymbol{Az}=\tilde{\boldsymbol{x}}=\boldsymbol{AA}^+\boldsymbol{y}$$

中的向量 $\boldsymbol{z}\in C^n$ 不是唯一的. 但是，根据(34)式，在满足 $\boldsymbol{Az}=\tilde{\boldsymbol{x}}$ 的向量中有最小范数的向量始终是唯一的，它可以由 $\hat{\boldsymbol{z}}=\boldsymbol{A}^+\boldsymbol{y}$ 给出.

(iii) 详细内容将在第II部分讲述. 这里只给出一些结果. 对于任意的 $\boldsymbol{y}\in C^m$, $\boldsymbol{z}\in C^n$，下列关系

$\boldsymbol{AA}^+\boldsymbol{y}=\boldsymbol{y}$ 到 $R(\boldsymbol{A})$ 的正交投影

$(\boldsymbol{I}-\boldsymbol{A}^+\boldsymbol{A})\boldsymbol{z}=\boldsymbol{z}$ 到 $N(\boldsymbol{A})$ 的正交投影

$\boldsymbol{A}^+\boldsymbol{Az}=\boldsymbol{z}$ 到 $R(\boldsymbol{A}^*)$ 的正交投影

$(\boldsymbol{I}-\boldsymbol{AA}^+)\boldsymbol{y}=\boldsymbol{y}$ 到 $N(\boldsymbol{A}^*)$ 的正交投影

成立.

在矩阵方程式 $\boldsymbol{AXB}=\boldsymbol{C}$ 中的应用

下面我们用伪逆矩阵来考察矩阵方程式 $\boldsymbol{AXB}=\boldsymbol{C}$ 的解，其中假定 $\boldsymbol{A}\in C^{m\times n}$, $\boldsymbol{B}\in C^{p\times q}$, $\boldsymbol{C}\in C^{m\times q}$. 这个所谓的解就是满足 $\boldsymbol{AXB}=\boldsymbol{C}$ 的矩阵 $\boldsymbol{X}\in C^{n\times p}$*. 显然，该方程式是向量 $\boldsymbol{x}$ 的线性方程组 $\boldsymbol{Ax}=\boldsymbol{y}$ 的扩展形. 特别是，在 $p=q=1$, $\boldsymbol{B}=1$ 的情况下就相当于 $\boldsymbol{Ax}=\boldsymbol{y}$.

[定理4]

(i) $\boldsymbol{AXB}=\boldsymbol{C}$ 有解 $\Leftrightarrow \boldsymbol{AA}^-\boldsymbol{CB}^-\boldsymbol{B}=\boldsymbol{C}$

(ii) $\boldsymbol{AXB}=\boldsymbol{C}$ 的一般解可由下式

$$\boldsymbol{X}=\boldsymbol{A}^-\boldsymbol{CB}^-+(\boldsymbol{Z}+\boldsymbol{A}^-\boldsymbol{AZBB}^-),\quad \boldsymbol{Z}\in C^{n\times p}(\text{任意矩阵}) \tag{35}$$

* 原文误为 $\boldsymbol{X}=C^{n\times p}$. ——译者注

给出.

(iii) 在 $\boldsymbol{AXB}=\boldsymbol{C}$ 的解中, $\hat{\boldsymbol{X}}=\boldsymbol{A}^{+}\boldsymbol{C}\boldsymbol{B}^{+}$ 有最小范数. 其中使用的矩阵范数是 $\|\boldsymbol{X}\|^2\triangleq \mathrm{tr}(\boldsymbol{X}^*\boldsymbol{X})$.

(iv) $\boldsymbol{AXB}=\boldsymbol{C}$ 无解的情况下, 在所有最小二乘近似解中, $\hat{\boldsymbol{X}}=\boldsymbol{A}^{+}\boldsymbol{C}\boldsymbol{B}^{+}$ 具有最小范数. 即 $\hat{\boldsymbol{X}}=\boldsymbol{A}^{+}\boldsymbol{C}\boldsymbol{B}^{+}$ 属于满足

$$\|\boldsymbol{AXB}-\boldsymbol{C}\|\leqslant\|\boldsymbol{AZB}-\boldsymbol{C}\|,\ {}^{\forall}\boldsymbol{Z}\in C^{n\times p} \tag{36}$$

的 $\boldsymbol{X}\in C^{n\times p}$ 的集合, 而且对于该集合中任意的 $\boldsymbol{X}$ 有下列性质

$$\|\hat{\boldsymbol{X}}\|\leqslant\|\boldsymbol{X}\|, \tag{37}$$

其中假定 $\|\boldsymbol{X}\|^2\triangleq \mathrm{tr}(\boldsymbol{X}^*\boldsymbol{X})$.

(证明)

(i) ($\Rightarrow$): $\boldsymbol{AXB}=\boldsymbol{C}\Rightarrow \boldsymbol{AA}^{-}\boldsymbol{CB}^{-}\boldsymbol{B}$

$$=\boldsymbol{AA}^{-}(\boldsymbol{AXB})\boldsymbol{B}^{-}\boldsymbol{B}=\boldsymbol{AXB}=\boldsymbol{C}.$$

($\Leftarrow$): $\boldsymbol{AA}^{-}\boldsymbol{CB}^{-}\boldsymbol{B}=\boldsymbol{C}\Rightarrow \boldsymbol{A}^{-}\boldsymbol{CB}^{-}$ 是解.

(ii) 因 $\boldsymbol{X}=\boldsymbol{A}^{-}\boldsymbol{CB}^{-}+(\boldsymbol{Z}-\boldsymbol{A}^{-}\boldsymbol{AZBB}^{-})$

$$\Rightarrow \boldsymbol{AXB}=\boldsymbol{AA}^{-}\boldsymbol{CB}^{-}\boldsymbol{B}+\boldsymbol{A}(\boldsymbol{Z}-\boldsymbol{A}^{-}\boldsymbol{AZBB}^{-})\boldsymbol{B}=\boldsymbol{C}$$

(利用(i)), 则 $\boldsymbol{X}$ 是解. 而且, 当 $\boldsymbol{Y}$ 为任意解时, 令 $\boldsymbol{Z}=\boldsymbol{Y}-\boldsymbol{A}^{-}\boldsymbol{CB}^{-}$, 则可以表示成

$$\boldsymbol{Y}=\boldsymbol{A}^{-}\boldsymbol{CB}^{-}+(\boldsymbol{Z}-\boldsymbol{A}^{-}\boldsymbol{AZBB}^{-})$$

(试证明).

(iii), (iv) 矩阵方程式 $\boldsymbol{AXB}=\boldsymbol{C}$ 不外乎是关于 $\boldsymbol{X}$ 的元素的线性方程组. 这个事实可以清楚地用克罗内克积(参看 B-I-4)表示.

现在考虑

$$\boldsymbol{A}=[a_{11}\ \ a_{12}],\ \boldsymbol{B}=\begin{bmatrix} b_{11} & b_{12}\\ b_{21} & b_{22}\end{bmatrix},\ \boldsymbol{C}=[C_{11}\ \ C_{12}],\ \boldsymbol{X}=\begin{bmatrix} x_{11} & x_{12}\\ x_{21} & x_{22}\end{bmatrix}$$

的情况, $\boldsymbol{AXB}=\boldsymbol{C}$ 可以表示成

$$\begin{bmatrix} a_{11}b_{11} & a_{12}b_{11} & a_{11}b_{21} & a_{12}b_{21}\\ a_{11}b_{12} & a_{12}b_{12} & a_{11}b_{22} & a_{12}b_{22}\end{bmatrix}\begin{bmatrix} x_{11}\\ x_{21}\\ x_{12}\\ x_{22}\end{bmatrix}=\begin{bmatrix} c_{11}\\ c_{12}\end{bmatrix}.$$

在这里, 若设将 $\boldsymbol{X}$ 及 $\boldsymbol{C}$ 的列依次排列成的向量, 即 $\boldsymbol{X}$, $\boldsymbol{C}$ 的列展开为 $\boldsymbol{x}$ 及 $\mathbf{c}$, 则

$$\boldsymbol{x}=[x_{11}x_{21}x_{12}x_{22}]^T,\ \mathbf{c}=[c_{11}c_{12}]^T,$$

上面关于 $\boldsymbol{X}$ 的元素的线性方程组可以写成

$$(\boldsymbol{B}^T\otimes\boldsymbol{A})\boldsymbol{x}=\mathbf{c}. \tag{38}$$

在一般情况下也完全一样, 将 $\boldsymbol{X}$ 及 $\boldsymbol{C}$ 的元素按上述原则排列成向量, 利用 B-I-4 中(18)式, 则 $\boldsymbol{AXB}=\boldsymbol{C}$ 可以表示成(38)式的形式.

将定理 1 用于(38)式, 则

$$\hat{\boldsymbol{x}}=(\boldsymbol{B}^T\otimes\boldsymbol{A})^{+}\mathbf{c} \tag{39}$$

是(38)式的最小二乘近似解中有最小范数的解. 即, 对于满足

$$\|(\boldsymbol{B}^T\otimes\boldsymbol{A})\boldsymbol{x}-\mathbf{c}\|^2\leqslant\|(\boldsymbol{B}^T\otimes\boldsymbol{A})\boldsymbol{z}-\mathbf{c}\|^2,\ {}^{\forall}\boldsymbol{z}\in C^{n\times p} \tag{40}$$

的向量 $\boldsymbol{x}$,

$$\|\hat{\boldsymbol{x}}\|^2 \leqslant \|\boldsymbol{x}\|^2 \tag{41}$$

成立. 但是,因这里所使用的向量范数是欧几里得范数,则

$$\|\boldsymbol{x}\|^2=(x_{11}^2+x_{12}^2+\cdots+x_{1n}^2)+\cdots+(x_{n1}^2+x_{n2}^2+\cdots+x_{nn}^2)=\operatorname{tr}(\boldsymbol{X}^*\boldsymbol{X}) \tag{42}$$

成立. 因此,若(40), (41)式按着 $\boldsymbol{x}\to\boldsymbol{X}$, $\boldsymbol{c}\to\boldsymbol{C}$ 改写,则分别变成

$$\|\boldsymbol{AXB}-\boldsymbol{C}\|^2\leqslant\|\boldsymbol{AZB}-\boldsymbol{C}\|^2,\ {}^\forall\boldsymbol{Z}\in C^{n\times p}, \tag{40$'$}$$

$$\|\hat{\boldsymbol{X}}\|^2\leqslant\|\boldsymbol{X}\|^2, \tag{41$'$}$$

式中 $\|\boldsymbol{X}\|^2=\operatorname{tr}(\boldsymbol{X}^*\boldsymbol{X})$. 最后, 因 $(\boldsymbol{B}^T\otimes\boldsymbol{A})^+=(\boldsymbol{B}^T)^+\otimes\boldsymbol{A}^+$, 再用 B-I-4 中(18)改写成矩阵,则 $\hat{\boldsymbol{x}}=(\boldsymbol{B}^T\otimes\boldsymbol{A})^+\boldsymbol{c}$ 变成

$$\hat{\boldsymbol{X}}=\boldsymbol{A}^+\boldsymbol{C}\boldsymbol{B}^+. \tag{43}$$

在 $\boldsymbol{AXB}=\boldsymbol{C}$ 有解的情况下, 由(40)$'$式, 对于满足 $\boldsymbol{AXB}=\boldsymbol{C}$ 的 $\boldsymbol{X}$, $\hat{\boldsymbol{X}}$ 是满足(41)$'$式的解.

证明完毕

A-I-16 离散时间系统

仅考虑在离散时刻 $t=\cdots t_{-2},\ t_{-1},\ t_0,\ t_1,\ t_2,\ \cdots$ 行为的动力学系统模型，可以用下列正规形差分方程式

$$\boldsymbol{x}(t_{k+1})=\boldsymbol{f}(t_k;\ \boldsymbol{x}(t_k);\ \boldsymbol{u}(t_k)) \tag{1}$$

$$\boldsymbol{y}(t_k)=\boldsymbol{r}(t_k;\ \boldsymbol{x}(t_k);\ \boldsymbol{u}(t_k)) \tag{2}$$

表示(参看 A-I-1). 而且，当上列(1)，(2)式右边是输入和状态的线性函数，即可以表示成

$$\boldsymbol{x}(t_{k+1})=\boldsymbol{F}(t_k)\boldsymbol{x}(t_k)+\boldsymbol{G}(t_k)\boldsymbol{u}(t_k)\quad \boldsymbol{x}(t_k)\in R^n,\ \boldsymbol{u}(t_k)\in \boldsymbol{R}^r \tag{3}$$

$$\boldsymbol{y}(t_k)=\boldsymbol{C}(t_k)\boldsymbol{x}(t_k)+\boldsymbol{D}(t_k)\boldsymbol{u}(t_k)\quad \boldsymbol{y}(t_k)\in R^m \tag{4}$$

时，称之为线性标准形差分方程式(参看 A-I-1).

以下为了简单起见，假定所考察时刻是等间隔的，即对于某个起始时刻 t_0 和某个 $T>0$ 可以写成

$$t_k=t_0+kT,\ k=0,\ \pm1,\ \pm2,\ \cdots, \tag{5}$$

$\boldsymbol{x}(t_0+kT)$ 等只略记成 $\boldsymbol{x}(k)$ 等等.

照例，我们从(3)式解的存在及唯一性开始讨论. 因(3)式右边仅仅是矩阵的乘法运算，若 $\boldsymbol{x}(k)$ 和 $\boldsymbol{u}(k)$ 给出，则 $\boldsymbol{x}(k+1)$ 唯一确定. 若 $\boldsymbol{x}(k+1)$ 和 $\boldsymbol{u}(k+1)$ 已确定，则 $\boldsymbol{x}(k+2)$ 唯一确定. 依此类推，当任意起始条件 $\boldsymbol{x}(k)$ 和任意输入序列 $\boldsymbol{u}(k),\ \boldsymbol{u}(k+1),\ \cdots$ 固定时，(3)式的解 $\boldsymbol{x}(l)\ \ l\geqslant k$ 存在而且唯一确定.

在连续时间动力学系统的情况下，若状态方程式满足利普希茨条件(参看 P-I-3)，则在 t_0 时刻通过 $\boldsymbol{x}_0$ 的解在 $t>t_0$ 时是(当然在过去的时间 $t<t_0$ 也是)唯一确定的(P-I-3 中注意(iii)). 在这里所考察的离散时间系统中，该性质一般不成立. 例如在 $\boldsymbol{F}(k-1)=\boldsymbol{0}$ 的情况下(3)式变为

$$\boldsymbol{x}(k)=\boldsymbol{0}\boldsymbol{x}(k-1)+\boldsymbol{G}(k-1)\boldsymbol{u}(k-1), \tag{6}$$

虽然给出 $\boldsymbol{x}(k)$ 和 $\boldsymbol{u}(k-1)$，但是 $\boldsymbol{x}(k-1)$ 不是唯一确定的. $\boldsymbol{x}(k)$ 和 $\boldsymbol{u}$ 已给出，即使对于过去时间($l\leqslant k$)解也唯一确定的充分必要条件是，$\boldsymbol{F}(k)$ 是正则的，即

$$\det \boldsymbol{F}(k)\neq 0\quad k=0,\ \pm1,\ \pm2,\ \cdots. \tag{7}$$

这里我们再稍讲一下，在那些情况下能得出(3)，(4)式表示的模型. 第1种情况是，系统由仅考虑原来离散时刻行为的环节(具有代表性的是滞后环节(delay))或子系统组成.

离散时间动力学系统标准形的推导方法(第1种情况)

这里以仅考虑原来离散时刻行为的环节或子系统组合成的系统为对象，讨论其标准形的推导方法.

现在假定系统的方框图由 P 个方框(子系统)构成，各方框的输入输出特性由下列标准形

$$\boldsymbol{x}_i(k+1)=\widetilde{\boldsymbol{F}}_i\boldsymbol{x}_i(k)+\widetilde{\boldsymbol{G}}_i\boldsymbol{u}_i(k) \tag{8a}$$

$$\boldsymbol{y}_i(k)=\boldsymbol{C}_i\boldsymbol{x}_i(k)+\boldsymbol{D}_i\boldsymbol{u}_i(k)\quad i=1,\ 2,\ \cdots,\ P \tag{8b}$$

给出. 注意，作为其特殊情况，也包含方框仅由延迟器($\widetilde{\boldsymbol{F}}_i=\boldsymbol{0}$, $\boldsymbol{D}_i=\boldsymbol{0}$)或放大器($\widetilde{\boldsymbol{F}}_i=\boldsymbol{0}$, $\widetilde{\boldsymbol{G}}_i=\boldsymbol{0}$, $\boldsymbol{C}_i=\boldsymbol{0}$)构成的场合.

如 A-I-4 中关于连续系统所述，将整个方框图的表示式变成所谓的中间标准形. 即，设

$$\boldsymbol{X}\triangleq[\boldsymbol{x}_1^T,\ \boldsymbol{x}_2^T,\ \cdots,\ \boldsymbol{x}_P^T]^T=\text{子系统的状态}$$

$$\boldsymbol{U}\triangleq[\boldsymbol{u}_1^T,\ \boldsymbol{u}_2^T,\ \cdots,\ \boldsymbol{u}_P^T]^T=\text{子系统的输入}$$

$$\boldsymbol{Y}\triangleq[\boldsymbol{y}_1^T,\ \boldsymbol{y}_2^T,\ \cdots,\ \boldsymbol{y}_P^T]^T=\text{子系统的输出}$$

$$\boldsymbol{V}\triangleq\text{整个系统的输入}$$

$$\boldsymbol{W}\triangleq\text{整个系统的输出,}$$

将子系统的特性归纳后写成

$$\boldsymbol{X}(k+1)=\widetilde{\boldsymbol{F}}\boldsymbol{X}(k)+\widetilde{\boldsymbol{G}}\boldsymbol{U}(k) \tag{9a}$$

$$\boldsymbol{Y}(k)=\boldsymbol{C}\boldsymbol{X}(k)+\boldsymbol{D}\boldsymbol{U}(k), \tag{9b}$$

式中

$$\widetilde{\boldsymbol{F}}\triangleq\mathrm{diag}[\widetilde{\boldsymbol{F}}_1,\ \widetilde{\boldsymbol{F}}_2,\ \cdots,\ \widetilde{\boldsymbol{F}}_P]$$

$$\widetilde{\boldsymbol{G}}\triangleq\mathrm{diag}[\widetilde{\boldsymbol{G}}_1,\ \widetilde{\boldsymbol{G}}_2,\ \cdots,\ \widetilde{\boldsymbol{G}}_P]$$

$$\boldsymbol{C}\triangleq\mathrm{diag}[\boldsymbol{C}_1,\ \boldsymbol{C}_2,\ \cdots,\ \boldsymbol{C}_P]$$

$$\boldsymbol{D}\triangleq\mathrm{diag}[\boldsymbol{D}_1,\ \boldsymbol{D}_2,\ \cdots,\ \boldsymbol{D}_P],$$

而且，仿照 A-I-4，将方框的结合用适当的结合矩阵 $\boldsymbol{F}$, $\boldsymbol{G}$, $\boldsymbol{J}$, $\boldsymbol{K}$ 表示成

$$\boldsymbol{U}=\boldsymbol{F}\boldsymbol{Y}+\boldsymbol{G}\boldsymbol{V} \tag{10}$$

$$\boldsymbol{W}=\boldsymbol{J}\boldsymbol{Y}+\boldsymbol{K}\boldsymbol{V}. \tag{11}$$

将(10)，(11)和(9)式归纳后可以得到和 A-I-4 中(9)式相同的中间标准形

$$\begin{bmatrix}\boldsymbol{I} & -\boldsymbol{F} & \boldsymbol{0} & \boldsymbol{0}\\ -\boldsymbol{D} & \boldsymbol{I} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & -\boldsymbol{J} & \boldsymbol{I} & \boldsymbol{0}\\ -\widetilde{\boldsymbol{G}} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}\end{bmatrix}\begin{bmatrix}\boldsymbol{U}(k)\\ \boldsymbol{Y}(k)\\ \boldsymbol{W}(k)\\ \boldsymbol{X}(k+1)\end{bmatrix}=\begin{bmatrix}\boldsymbol{0}\\ \boldsymbol{C}\\ \boldsymbol{0}\\ \widetilde{\boldsymbol{F}}\end{bmatrix}\boldsymbol{X}(k)+\begin{bmatrix}\boldsymbol{G}\\ \boldsymbol{0}\\ \boldsymbol{K}\\ \boldsymbol{0}\end{bmatrix}\boldsymbol{V}(k). \tag{12}$$

当且仅当 $\det[\boldsymbol{I}-\boldsymbol{D}\boldsymbol{F}]\ (=\det[\boldsymbol{I}-\boldsymbol{F}\boldsymbol{D}])\neq 0$ 时，(12)式左边的系数矩阵是正则的. 若 $\det[\boldsymbol{I}-\boldsymbol{D}\boldsymbol{F}]\neq 0$，则和 A-I-4 中(11)式一样，表示整个系统的标准形为

$$\boldsymbol{X}(k+1)=\{\widetilde{\boldsymbol{F}}+\widetilde{\boldsymbol{G}}(\boldsymbol{I}-\boldsymbol{F}\boldsymbol{D})^{-1}\boldsymbol{F}\boldsymbol{C}\}\boldsymbol{X}(k)+\widetilde{\boldsymbol{G}}(\boldsymbol{I}-\boldsymbol{F}\boldsymbol{D})^{-1}\boldsymbol{G}\boldsymbol{V}(k) \tag{13a}$$

$$\boldsymbol{W}(k)=\boldsymbol{J}(\boldsymbol{I}-\boldsymbol{D}\boldsymbol{F})^{-1}\boldsymbol{C}\boldsymbol{X}(k)+\{\boldsymbol{K}+\boldsymbol{J}(\boldsymbol{I}-\boldsymbol{D}\boldsymbol{F})^{-1}\boldsymbol{D}\boldsymbol{G}\}\boldsymbol{V}(k). \tag{13b}$$

根据(12)式，$\boldsymbol{U}(k)$, $\boldsymbol{Y}(k)$ 可用 $\boldsymbol{X}(k)$, $\boldsymbol{V}(k)$ 唯一表示. 当 $\boldsymbol{X}(0)$ 和输入序列 $\{\boldsymbol{V}(k)\}$ 给出时，因 $\boldsymbol{X}(k)$ 具有由(13a)式规定的唯一解，于是 $\boldsymbol{X}(k)$, $\boldsymbol{U}(k)$, $\boldsymbol{Y}(k)$, $\boldsymbol{W}(k)$ 都具有由 $\boldsymbol{X}(0)$ 和 $\{\boldsymbol{V}(k)\}$ 规定的唯一解.

若 $\det[\boldsymbol{I}-\boldsymbol{D}\boldsymbol{F}]\neq 0$，在(12)式的上半部分

$$\begin{bmatrix}\boldsymbol{I} & -\boldsymbol{F}\\ -\boldsymbol{D} & \boldsymbol{I}\end{bmatrix}\begin{bmatrix}\boldsymbol{U}(k)\\ \boldsymbol{Y}(k)\end{bmatrix}=\begin{bmatrix}\boldsymbol{0}\\ \boldsymbol{C}\end{bmatrix}\boldsymbol{X}(k)+\begin{bmatrix}\boldsymbol{G}\\ \boldsymbol{0}\end{bmatrix}\boldsymbol{V}(k) \tag{14}$$

中，因左边系数矩阵是非正则的，则进行适当的基本行变换可使零行向量出现在最下行，即

$$\begin{bmatrix} \tilde{\boldsymbol{S}}_1 \\ \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{U}(k) \\ \boldsymbol{Y}(k) \end{bmatrix} = \begin{bmatrix} \tilde{\boldsymbol{C}}_1 \\ \tilde{\boldsymbol{c}}_2 \end{bmatrix} \boldsymbol{X}(k) + \begin{bmatrix} \tilde{\boldsymbol{G}}_1 \\ \tilde{\boldsymbol{g}}_2 \end{bmatrix} \boldsymbol{V}(k). \tag{15}$$

(15)式的最下行是

$$\tilde{\boldsymbol{c}}_2 \boldsymbol{X}(k) + \tilde{\boldsymbol{g}}_2 \boldsymbol{V}(k) = \boldsymbol{0}, \tag{16}$$

(i)若 $\tilde{\boldsymbol{c}}_2 \neq \boldsymbol{0}$, 则(16)式意味着 $\boldsymbol{X}(k)$ 的各分量不是独立的，整个系统的状态向量可以用比 $\boldsymbol{X}$ 的维数小的向量表示. (ii)若 $\tilde{\boldsymbol{c}}_2 = \boldsymbol{0}$, $\tilde{\boldsymbol{g}}_2 \neq 0$, 则输入向量的各分量不是独立的，就是说只加上满足某种从属关系的输入，我们不考虑这种情况. (iii)若 $\tilde{\boldsymbol{c}}_2 = \boldsymbol{0}$, $\tilde{\boldsymbol{g}}_2 = \boldsymbol{0}$, 则 $\boldsymbol{U}(k)$, $\boldsymbol{Y}(k)$ 不能由 $\boldsymbol{X}(k)$, $\boldsymbol{V}(k)$ 唯一确定. 因此，虽然给出 $\boldsymbol{X}(0)$, $\{\boldsymbol{V}(k)\}$, 但表示整个系统的(12)式并不具有唯一解. 将以上结果归纳如下.

"具有(9)式特性的子系统按(10), (11)式结合起来构成整个系统，考虑这种离散时间系统，这时整个系统可以以各子系统状态的直和 $\boldsymbol{X} \triangleq [\boldsymbol{x}_1^T, \boldsymbol{x}_2^T, \cdots, \boldsymbol{x}_P^T]^T$ 为状态，用关于 $\boldsymbol{X}$ 的标准形表示. 而且，子系统输入 $\boldsymbol{U}(k)$ 和子系统输出 $\boldsymbol{Y}(k)$ 由 $\boldsymbol{X}(0)$ 和$\{\boldsymbol{V}(k)\}$唯一确定的充分必要条件是，$\det[\boldsymbol{I}-\boldsymbol{DF}]$ $(=\det[\boldsymbol{I}-\boldsymbol{FD}]) \neq 0$. 其中假定 $\boldsymbol{V}(k)$ 的各分量在各时点取独立值."

(注意)

(i) 在子系统具有变系数矩阵 $\tilde{\boldsymbol{F}}(k)$, $\tilde{\boldsymbol{G}}(k)$, $\boldsymbol{C}(k)$, $\boldsymbol{D}(k)$ 的情况下，当且仅当 $\det[\boldsymbol{I}-\boldsymbol{FD}(k)]$ $(=\det[\boldsymbol{I}-\boldsymbol{D}(k)\boldsymbol{F}]) \neq 0$, $^\forall k$ 时，可以得到和上面同样的结论.

(ii) $\det[\boldsymbol{I}-\boldsymbol{DF}] \neq 0$ 是整个系统以 $\boldsymbol{X}=[\boldsymbol{x}_1^T, \boldsymbol{x}_2^T, \cdots, \boldsymbol{x}_P^T]$ 为最小维数状态向量的充分必要条件. 而且也有这样一种情况，即使 $\det[\boldsymbol{I}-\boldsymbol{DF}]=0$, 在某种条件下，对于满足 $\dim \tilde{\boldsymbol{X}} < \dim \boldsymbol{X}$ 的适当的向量 $\tilde{\boldsymbol{X}}$, 整个系统可以用标准形表示，同时 $\boldsymbol{U}(k)$, $\boldsymbol{Y}(k)$ 也具有由 $\tilde{\boldsymbol{X}}(0)$, $\{\boldsymbol{V}(k)\}$ 规定的唯一解.

关于这方面的详细内容可参考文献[131].

(第2种情况)

第2种情况是，对连续时间系统的状态进行采样，仅在离散时刻取值的场合

在连续时间线性系统

$$\dot{\boldsymbol{x}} = \boldsymbol{A}(t)\boldsymbol{x} + \boldsymbol{B}(t)\boldsymbol{u} \quad \boldsymbol{x}(t_0) = \boldsymbol{x}_0 \tag{17a}$$

$$\boldsymbol{y} = \boldsymbol{C}(t)\boldsymbol{x} + \boldsymbol{D}(t)\boldsymbol{u} \tag{17b}$$

中，假定在时间间隔 $t_k \leqslant t < t_{k+1}$ 上输入一定，即

$$\boldsymbol{u}(t) = \boldsymbol{u}(k) \quad t_k \leqslant t < t_{k+1} \quad k=0, \pm 1, \pm 2, \cdots$$

考虑到该条件，对于从 t_k 到 t_{k+1} 将(17a)式解出(参照 A-I-6 中定理6)，得

$$\begin{aligned} \boldsymbol{x}(k+1) &= \boldsymbol{\Phi}(\overline{k+1}\ T, kT)\boldsymbol{x}(k) \\ &\quad + \int_{kT}^{\overline{k+1}T} \boldsymbol{\Phi}(\overline{k+1}\ T, t)\boldsymbol{B}(t)dt\ \boldsymbol{u}(k), \quad k=0, \pm 1, \pm 2, \cdots, \end{aligned} \tag{18}$$

式中 $\boldsymbol{\Phi}(\cdot, \cdot)$ 是自由系统

$$\dot{\boldsymbol{x}} = \boldsymbol{A}(t)\boldsymbol{x}$$

的状态转移矩阵. 对比(18)和(3)式可见，在这种情况下

$$\boldsymbol{F}(k) = \boldsymbol{\Phi}(\overline{k+1}\ T, kT) \tag{19}$$

$$\boldsymbol{G}(k)=\int_{kT}^{\overline{k+1}\,T}\boldsymbol{\Phi}(\overline{k+1}\ T,\ t)\boldsymbol{B}(t)dt. \tag{20}$$

因为状态转移矩阵是正则的(参照 A-I-6 定理 5)，则(19)式的 $\boldsymbol{F}(k)$ 满足(7)式. 这是第 2 种情况的特点. 在对连续时间系统采样得到的离散时间系统中，不论是什么样的系统，什么样的采样间隔 T，(7)式始终成立.

在(17)式中若只考察 t_k 时刻，显然可以得到(4)式.

现在我们来讨论连续时间系统是常系数的情况. 这种情况下状态转移矩阵 $\boldsymbol{\Phi}(t,\ s)$ 可以给成(参照 A-I-9)

$$\boldsymbol{\Phi}(t,\ s)=e^{A(t-s)}.$$

将其代入(19)式，得

$$\boldsymbol{F}(k)=e^{A(\overline{k+1}T-kT)}=e^{AT}\triangleq\boldsymbol{F} \tag{21}$$

再代入(20)式，考虑到 $\boldsymbol{B}$ 是常数矩阵，则

$$\boldsymbol{G}(k)=\int_{kT}^{\overline{k+1}\ T}e^{A(\overline{k+1}\ T-t)}dt\boldsymbol{B}.$$

设 $\tau\triangleq\overline{k+1}\,T-t$，进行变量变换后，得

$$\boldsymbol{G}(k)=\int_{T}^{0}e^{A\tau}(-d\tau)\boldsymbol{B}=\int_{0}^{T}e^{A\tau}\,d\tau\,\boldsymbol{B}\triangleq\boldsymbol{G}, \tag{22}$$

都变成与 k 无关的常数矩阵(和预想的一样).

状态方程式的解

实际上写出(3)式解是很容易的，一般 $\boldsymbol{x}(k)\ (k\geqslant 0)$ 可给成

$$\begin{aligned}\boldsymbol{x}(k)=&\boldsymbol{F}(k-1)\cdots\boldsymbol{F}(1)\boldsymbol{F}(0)\boldsymbol{x}(0)+\boldsymbol{G}(k-1)\boldsymbol{u}(k-1)\\&+\boldsymbol{F}(k-1)\boldsymbol{G}(k-2)\boldsymbol{u}(k-2)+\cdots+\boldsymbol{F}(k-1)\times\cdots\times\boldsymbol{F}(1)\boldsymbol{G}(0)\boldsymbol{u}(0).\end{aligned}$$

为了简单起见，也常写成

$$\boldsymbol{x}(k)=\prod_{j=0}^{k-1}\boldsymbol{F}(j)\boldsymbol{x}(0)+\sum_{i=0}^{k-1}\left\{\prod_{j=i+1}^{k-1}\boldsymbol{F}(j)\right\}\boldsymbol{G}(i)\boldsymbol{u}(i). \tag{23}$$

因式中 $\prod$ 表示矩阵的乘法运算，则必须特别注意乘法运算的顺序.

将(23)式代入(4)式可以立即写出输出

$$\boldsymbol{y}(k)=\boldsymbol{C}(k)\prod_{j=0}^{k-1}\boldsymbol{F}(j)\boldsymbol{x}(0)+\boldsymbol{C}(k)\sum_{i=0}^{k-1}\left\{\prod_{j=i+1}^{k-1}\boldsymbol{F}(j)\right\}\boldsymbol{G}(i)\boldsymbol{u}(i)+\boldsymbol{D}(k)\boldsymbol{u}(k). \tag{24}$$

常系数情况下更简单一些，

$$\boldsymbol{x}(k)=\boldsymbol{F}^{k}\boldsymbol{x}(0)+\sum_{i=0}^{k-1}\boldsymbol{F}^{k-i-1}\boldsymbol{G}\boldsymbol{u}(i) \tag{25}$$

$$\boldsymbol{y}(k)=\boldsymbol{C}\boldsymbol{F}^{k}\boldsymbol{x}(0)+\boldsymbol{C}\sum_{i=0}^{k-1}\boldsymbol{F}^{k-i-1}\boldsymbol{G}\boldsymbol{u}(i)+\boldsymbol{D}\boldsymbol{u}(k). \tag{26}$$

此外，将向量 $\boldsymbol{x}(k)$ 的 Z 变换 $\hat{\boldsymbol{x}}(Z)$ 定义成以各分量的 Z 变换为分量的向量，则

$$\hat{\boldsymbol{x}}(Z)\triangleq\boldsymbol{x}(0)+Z^{-1}\boldsymbol{x}(1)+Z^{-2}\boldsymbol{x}(2)+\cdots+Z^{-k}\boldsymbol{x}(k)+\cdots.$$

对于 $\boldsymbol{u}(k)$，$\boldsymbol{y}(k)$ 也同样，由常系数系统的状态方程式立即可得

$$
\begin{aligned}
&\boldsymbol{x}(1)+Z^{-1}\boldsymbol{x}(2)+\cdots+Z^{-k}\boldsymbol{x}(k+1)+\cdots \\
&\quad=\boldsymbol{F}[\boldsymbol{x}(0)+Z^{-1}\boldsymbol{x}(1)+\cdots+Z^{-k}\boldsymbol{x}(k)+\cdots] \\
&\qquad+\boldsymbol{G}[\boldsymbol{u}(0)+Z^{-1}\boldsymbol{u}(1)+\cdots+Z^{-k}\boldsymbol{u}(k)+\cdots] \\
&\quad=\boldsymbol{F}\hat{\boldsymbol{x}}(Z)+\boldsymbol{G}\hat{\boldsymbol{u}}(Z).
\end{aligned}
$$

容易看到,上式左边等于 $Z\hat{\boldsymbol{x}}(Z)-Z\boldsymbol{x}(0)$,因此

$$
(Z\boldsymbol{I}-\boldsymbol{F})\hat{\boldsymbol{x}}(Z)=Z\boldsymbol{x}(0)+\boldsymbol{G}\hat{\boldsymbol{u}}(Z)
$$

$$
\therefore\quad \hat{\boldsymbol{x}}(Z)=(Z\boldsymbol{I}-\boldsymbol{F})^{-1}Z\boldsymbol{x}(0)+(Z\boldsymbol{I}-\boldsymbol{F})^{-1}\boldsymbol{G}\hat{\boldsymbol{u}}(Z). \tag{27}
$$

特别是在 $\boldsymbol{x}(0)=\boldsymbol{0}$ 的情况下

$$
\hat{\boldsymbol{x}}(Z)=(Z\boldsymbol{I}-\boldsymbol{F})^{-1}\boldsymbol{G}\hat{\boldsymbol{u}}(Z). \tag{28}
$$

而且,对于输出可得

$$
\hat{\boldsymbol{y}}(Z)=[\boldsymbol{C}(Z\boldsymbol{I}-\boldsymbol{F})^{-1}\boldsymbol{G}+\boldsymbol{D}]\hat{\boldsymbol{u}}(Z). \tag{29}
$$

(29)式右边[ ]内部分称为系统的脉冲传递函数.

可控性

现来考察系统(3)的可控性. 对于任意给定的 $\boldsymbol{x}_0\in R^n$ 和 $\boldsymbol{x}_1\in R^n$, 当存在某个自然数 k 和输入序列 $\boldsymbol{u}(0), \boldsymbol{u}(1), \cdots, \boldsymbol{u}(k)$, 使得

$$
\boldsymbol{x}(0)=\boldsymbol{x}_0 \tag{30}
$$

$$
\boldsymbol{x}(k+1)=\boldsymbol{x}_1 \tag{31}
$$

成立时,称为系统(3)在时刻 t_0 是可控的.

(23)式稍微变化一下,把(30), (31)式代入,得下列线性方程组

$$
\left[\boldsymbol{G}(k)\vdots\boldsymbol{F}(k)\boldsymbol{G}(k-1)\vdots\cdots\vdots\prod_{j=1}^{k}\boldsymbol{F}(j)\boldsymbol{G}(0)\right]\begin{bmatrix}\boldsymbol{u}(k)\\ \boldsymbol{u}(k-1)\\ \vdots\\ \boldsymbol{u}(0)\end{bmatrix}=\boldsymbol{x}_1-\prod_{j=0}^{k}\boldsymbol{F}(j)\boldsymbol{x}_0. \tag{32}
$$

若对于某个 k 左边系数矩阵的秩等于 n, 则对于任意 $\boldsymbol{x}_0, \boldsymbol{x}_1\in R^n$, 有满足该方程式的解 $\boldsymbol{u}(0), \cdots, \boldsymbol{u}(k)$ 存在,因此是可控的. 反之,若系数矩阵的秩小于 n, 因其行线性相关,则存在某个 $\boldsymbol{\eta}\in R^n$, $\boldsymbol{\eta}\neq\boldsymbol{0}$, 使得

$$
\boldsymbol{\eta}^T\left[\boldsymbol{G}(k)\vdots\boldsymbol{F}(k)\boldsymbol{G}(k-1)\vdots\cdots\vdots\prod_{j=1}^{k}\boldsymbol{F}(j)\boldsymbol{G}(0)\right]=\boldsymbol{0} \tag{33}
$$

成立. 若选取 $\boldsymbol{x}_0=\boldsymbol{0}$, $\boldsymbol{x}_1=\boldsymbol{\eta}$; 则不存在满足(32)式的 $\boldsymbol{u}(0), \cdots, \boldsymbol{u}(k)$. 其原因是,若存在的话,在(32)式两边左乘以 $\boldsymbol{\eta}^T$, 得

$$
\boldsymbol{0}=\boldsymbol{\eta}^T\boldsymbol{\eta}=\|\boldsymbol{\eta}\|_2^2,
$$

与 $\boldsymbol{\eta}\neq\boldsymbol{0}$ 相矛盾.

由以上得下列结果.

[定理 1] (3)式系统在时刻 t_0 可控⇔对于某个 k,

$$
\operatorname{rank}\left[\boldsymbol{G}(k)\vdots\boldsymbol{F}(k)\boldsymbol{G}(k-1)\vdots\cdots\vdots\prod_{j=1}^{k}\boldsymbol{F}(j)\boldsymbol{G}(0)\right]=n \tag{34}
$$

成立.

(问题 1) 试推导出和(34)式等价的按照克兰姆行列式判定的条件(指示:参考 B-I-9 中问题 5).

根据定理1和凯莱-哈密顿定理，关于常系数系统的可控性可得下列结果.

[系1] 常系数系统

$$\boldsymbol{x}(k+1)=\boldsymbol{F}\boldsymbol{x}(k)+\boldsymbol{G}\boldsymbol{u}(k)$$

可控

$$\Leftrightarrow \operatorname{rank}[\boldsymbol{G} \vdots \boldsymbol{F}\boldsymbol{G} \vdots \cdots \vdots \boldsymbol{F}^{n-1}\boldsymbol{G}]=n \tag{35}$$

(问题2) 试证明系1.

在连续时间系统中，由任意 $\boldsymbol{x}_0\in R^n$ 移到原点 $(\boldsymbol{x}_1=\boldsymbol{0})$ 的控制存在，和从任意 $\boldsymbol{x}_0\in R^n$ 移到任意 $\boldsymbol{x}_1\in R^n$ 的控制存在是等价的(参照 A-I-8). 在离散时间系统中该结果未必成立. 例如，在常系数系统中若设 $\boldsymbol{F}=\boldsymbol{0}$, $\boldsymbol{G}=\boldsymbol{0}$, 由(35)式显而易见，系统不是可控的. 尽管如此，因 $\boldsymbol{x}(1)=\boldsymbol{0}$, 则可从任意 $\boldsymbol{x}_0\in R^n$ 移到原点.

(问题3) 试证明"由任意 $\boldsymbol{x}_0$ 到原点可控

$\Leftrightarrow$对于某个 k,

$$\operatorname{rank}\left[\boldsymbol{G}(k) \vdots \boldsymbol{F}(k)\boldsymbol{G}(k-1) \vdots \cdots \vdots \prod_{j=1}^{k}\boldsymbol{F}(j)\boldsymbol{G}(0)\right]$$

$$=\operatorname{rank}\left[\boldsymbol{G}(k) \vdots \boldsymbol{F}(k)\boldsymbol{G}(k-1) \vdots \cdots \vdots \prod_{j=1}^{k}\boldsymbol{F}(j)\boldsymbol{G}(0) \vdots \prod_{j=0}^{k}\boldsymbol{F}(j)\right]"$$

(问题4) 对于(7)式成立的系统(和连续系统的情况一样)，试证明到原点的可控性和可控性是等价的.

可观测性

下面来讨论系统(3)，(4)的可观测性. 可观测性的定义和连续系统的情况(A-I-8)一样. 设 $\boldsymbol{u}(k)\equiv 0$, $k=0, 1, 2, \cdots$, 利用(24)式依次写出 $\boldsymbol{y}(0)$, $\boldsymbol{y}(1)$, $\cdots$, $\boldsymbol{y}(k)$, 得

$$\begin{aligned}
\boldsymbol{y}(0)&=\boldsymbol{C}(0)\boldsymbol{x}(0)\\
\boldsymbol{y}(1)&=\boldsymbol{C}(1)\boldsymbol{F}(0)\boldsymbol{x}(0)\\
&\ \vdots\\
\boldsymbol{y}(k)&=\boldsymbol{C}(k)\prod_{j=0}^{k-1}\boldsymbol{F}(j)\boldsymbol{x}(0)
\end{aligned}$$

即可得到下列线性方程组

$$\begin{bmatrix}\boldsymbol{C}(0)\\ \boldsymbol{C}(1)\boldsymbol{F}(0)\\ \vdots\\ \boldsymbol{C}(k)\prod\limits_{j=1}^{k-1}\boldsymbol{F}(j)\end{bmatrix}\boldsymbol{x}_0=\begin{bmatrix}\boldsymbol{y}(0)\\ \boldsymbol{y}(1)\\ \vdots\\ \boldsymbol{y}(k)\end{bmatrix}. \tag{36}$$

设左边系数矩阵为 $\boldsymbol{R}_k$, 对于某个 k, 若 $\operatorname{rank}\boldsymbol{R}_k=n$, 因 $\boldsymbol{R}_k^T\boldsymbol{R}_k$ 是正则的，根据

$$\boldsymbol{x}_0=(\boldsymbol{R}_k^T\boldsymbol{R}_k)^{-1}\boldsymbol{R}_k^T\begin{bmatrix}\boldsymbol{y}(0)\\ \boldsymbol{y}(1)\\ \vdots\\ \boldsymbol{y}(k)\end{bmatrix}, \tag{37}$$

则 $\boldsymbol{x}_0$ 由 $\boldsymbol{y}(0)$, $\boldsymbol{y}(1)$, $\cdots$, $\boldsymbol{y}(k)$ 唯一确定. 假定对于任何 k 都是 $\operatorname{rank}\boldsymbol{R}_k<n$, 则有满足 $\boldsymbol{R}_k\boldsymbol{x}_0^*=\boldsymbol{0}$ 的 $\boldsymbol{x}_0^*\in R^n$, $\boldsymbol{x}_0^*\neq\boldsymbol{0}$ 存在，即对于起始值 $\boldsymbol{x}_0^*$ 输出恒等于 $\boldsymbol{0}$. 因此，不能区别 $\boldsymbol{x}_0^*$ 和

0 两个起始值,不是可观测的. 于是得下列结果.

[定理 2] 系统(3), (4) 可观测

⇔对于某个 k,

$$\operatorname{rank}\begin{bmatrix} \boldsymbol{C}(0) \\ \boldsymbol{C}(1)\boldsymbol{F}(0) \\ \vdots \\ \boldsymbol{C}(k)\prod_{j=0}^{k-1}\boldsymbol{F}(j) \end{bmatrix}=n \tag{38}$$

成立.

[系 2] 常系数系统

$$\boldsymbol{x}(k+1)=\boldsymbol{F}\boldsymbol{x}(k)+\boldsymbol{G}\boldsymbol{u}(k)$$
$$\boldsymbol{y}(k)=\boldsymbol{C}\boldsymbol{x}(k)+\boldsymbol{D}\boldsymbol{u}(k)$$

可观测

⇔

$$\operatorname{rank}\begin{bmatrix} \boldsymbol{C} \\ \boldsymbol{C}\boldsymbol{F} \\ \vdots \\ \boldsymbol{C}\boldsymbol{F}^{n-1} \end{bmatrix}=n \tag{39}$$

在以上讨论中, 所谓可观测性定义成观测 $\boldsymbol{y}(0)$, …, $\boldsymbol{y}(k)$, 可以唯一确定 $\boldsymbol{x}(0)$, 这和观测 $\boldsymbol{y}(-k)$, …, $\boldsymbol{y}(-1)$, $\boldsymbol{y}(0)$ 可以唯一确定 $\boldsymbol{x}(0)$ 不是等价的. 例如, 在常系数 $\boldsymbol{F}=\boldsymbol{0}$ 的情况下, 因为对于任意 $\boldsymbol{x}(-1)$ 都是 $\boldsymbol{x}(0)=\boldsymbol{0}$, 则对于任何输出(既使输出不完全清楚)都可以唯一地求出 $\boldsymbol{x}(0)$. 但是,例如若 $\boldsymbol{C}=\boldsymbol{0}$, 则不是可观测的. 下面结果的证明, 利用第 II 部分将要讲的几何方法是很简单的,故这里证明省略.

"对于某个 $k>0$, 观测 $\boldsymbol{y}(-k)$, …, $\boldsymbol{y}(-1)$, $\boldsymbol{y}(0)$ 可以唯一地求出 $\boldsymbol{x}(0)$⇔对于某个 $k>0$."

$$\operatorname{rank}\begin{bmatrix} \boldsymbol{C}(-k) \\ \boldsymbol{C}(-k+1)\boldsymbol{F}(-k) \\ \vdots \\ \boldsymbol{C}(0)\prod_{j=1}^{k}\boldsymbol{F}(-j) \end{bmatrix}=\operatorname{rank}\begin{bmatrix} \boldsymbol{C}(-k) \\ \boldsymbol{C}(-k+1)\boldsymbol{F}(-k) \\ \vdots \\ \boldsymbol{C}(0)\prod_{j=1}^{k}\boldsymbol{F}(-j) \\ \prod_{j=0}^{k}\boldsymbol{F}(-j) \end{bmatrix}. \tag{40}$$

和连续时间系统的可控性(可观测性)的对比

在这里我们将常系数连续时间系统(S_c)

$$\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{u}$$
$$\boldsymbol{y}=\boldsymbol{C}\boldsymbol{x}+\boldsymbol{D}\boldsymbol{u}$$

和常系数离散时间系统(S_d)

$$\boldsymbol{x}(k+1)=\boldsymbol{F}\boldsymbol{x}(k)+\boldsymbol{G}\boldsymbol{u}(k)$$
$$\boldsymbol{y}(k)=\tilde{\boldsymbol{C}}\boldsymbol{x}(k)+\tilde{\boldsymbol{D}}\boldsymbol{u}(k)$$

的可控性(可观测性)进行对比.

首先注意，根据(35)，(39)式，为使 S_c 可控(可观测) $\boldsymbol{A}$，$\boldsymbol{B}$($\boldsymbol{C}$，$\boldsymbol{A}$) 应该满足的条件，和为使 S_d 可控 $\boldsymbol{F}$，$\boldsymbol{G}$($\widetilde{\boldsymbol{C}}$，$\boldsymbol{F}$) 应该满足的条件具有完全相同的形式. 因此，关于可控 S_c 的 $\boldsymbol{A}$，$\boldsymbol{B}$ 成立的性质，例如，为使 $\boldsymbol{A}$ 变换成若当形时可控 $\boldsymbol{B}$ 应该满足的条件，$\boldsymbol{A}$，$\boldsymbol{B}$ 的典范形，状态反馈的可控性保存，极点配置等结果，对于可控 S_d 的 $\boldsymbol{F}$，$\boldsymbol{G}$ 均成立. 对于可观测性也是同样.

其次，考虑 S_d 是由对 S_c 采样得到的，即

$$\boldsymbol{F}=e^{\boldsymbol{A}T},\ \boldsymbol{G}=\int_0^T e^{\boldsymbol{A}\tau}\,d\tau\,\boldsymbol{B} \tag{41}$$

$$\widetilde{\boldsymbol{C}}=\boldsymbol{C},\ \widetilde{\boldsymbol{D}}=\boldsymbol{D} \tag{42}$$

成立的情况.

[性质 1] 当 S_c 和 S_d 之间有上面(41)，(42)式的关系时，若 S_d 可控(可观测)，则 S_c 亦可控(可观测).

(问题 5) 试证明性质 1.

该性质的逆一般不成立. 设 $\boldsymbol{A}$ 的若当形为 $\boldsymbol{J}$，则 $e^{\boldsymbol{J}t}$ 可用 B-I-13 中(16a)～(16d)式给出. 由此及 B-I-11 中系 5，当设 λ_1，$\cdots$，λ_σ 为 $\boldsymbol{A}$ 的相异特征值时，则 $\boldsymbol{F}=e^{\boldsymbol{A}T}$ 的特征值为 $e^{\lambda_1 T}$，$\cdots$，$e^{\lambda_\sigma T}$.

这里假定 $\boldsymbol{A}$ 的某两个特征值，例如 λ_1 和 λ_2 之间有下列关系

$$\lambda_1-\lambda_2=j\,\frac{2\pi k}{T},\ k\text{ 是整数},\ k\neq 0,$$

则

$$e^{\lambda_1 T}=e^{\lambda_2 T+j2\pi k}=e^{\lambda_2 T},$$

尽管 $\boldsymbol{A}$ 的特征值 λ_1 和 λ_2 相异，但对应的 $\boldsymbol{F}$ 的特征值 $e^{\lambda_1 T}$ 和 $e^{\lambda_2 T}$ 相等. 这就是性质 1 的逆一般不成立的道理.

[性质 2] 假定 S_c 和 S_d 有(41)，(42)式的关系，S_c 可控(可观测). 这时，S_d 也可控(可观测)的充分条件是，对于任意自然数 k 和 p，q，$1\leqslant p$，$q\leqslant\sigma$，$p\neq q$，下式

$$\lambda_p-\lambda_q\neq j\,\frac{2\pi k}{T} \tag{43}$$

成立.

(证明) 先来证明在(41)式条件下

$$\boldsymbol{\eta}^T\boldsymbol{A}=\lambda\boldsymbol{\eta}^T,\ \boldsymbol{\eta}^T\boldsymbol{B}=\boldsymbol{0}\Leftrightarrow\boldsymbol{\eta}^T\boldsymbol{F}=e^{\lambda T}\boldsymbol{\eta}^T,\ \boldsymbol{\eta}^T\boldsymbol{G}=\boldsymbol{0}. \tag{44}$$

由 B-I-13 中(12)，(16)式，显然

$$\boldsymbol{\eta}^T\boldsymbol{A}=\lambda\boldsymbol{\eta}^T\Leftrightarrow\boldsymbol{\eta}^T e^{\boldsymbol{A}t}=e^{\lambda t}\boldsymbol{\eta}^T.$$

对于满足 $\boldsymbol{\eta}^T\boldsymbol{A}=\lambda\boldsymbol{\eta}^T$ 的任意 $\boldsymbol{\eta}\in C^n$，

$$\boldsymbol{\eta}^T\boldsymbol{G}=\boldsymbol{\eta}^T\int_0^T e^{\boldsymbol{A}\tau}\,d\tau\,\boldsymbol{B}=\int_0^T\boldsymbol{\eta}^T e^{\boldsymbol{A}\tau}\,d\tau\,\boldsymbol{B}=\int_0^T e^{\lambda\tau}\,d\tau\,\boldsymbol{\eta}^T\boldsymbol{B}$$

$$\therefore\quad \boldsymbol{\eta}^T\boldsymbol{B}=0\Leftrightarrow\boldsymbol{\eta}^T\boldsymbol{G}=0,$$

因此(44)式得证.

若假定 S_d 不是可控的，根据本章附录中的引理，有满足 $\boldsymbol{\eta}^T\boldsymbol{F}=\mu\boldsymbol{\eta}^T$，$\boldsymbol{\eta}^T\boldsymbol{G}=\boldsymbol{0}$ 的 $\boldsymbol{\eta}\neq\boldsymbol{0}$

存在. 由(44)式, $\boldsymbol{\eta}^T\boldsymbol{A}=\lambda\boldsymbol{\eta}^T$, $\boldsymbol{\eta}^T\boldsymbol{B}=\mathbf{0}$, 式中 $\lambda=\dfrac{1}{T}\ln\mu$. 再由附录中引理, S_c 不是可控的.

对于可观测性也是同样.

证明完毕

[性质 3] 在单输入系统($r=1$)的情况下, 根据(41), (42)式由可控 S_c 构成的 S_d 可控的充分必要条件是(43)式成立. 对于单输出系统($m=1$)的可观测性也是同样.

(证明) 由性质 2, 充分性是显然的.

例如, 对于 λ_1 和 λ_2, 假定(43)式不成立, 如上所述, $e^{\lambda_1 T}=e^{\lambda_2 T}$. 因此, $\boldsymbol{F}$ 的特征值 $e^{\lambda_1 T}$ 的几何重复度至少是 2. 根据 B-I-12 中系 1 和 A-I-11 中系 5, 单输入系统 S_d 不是可控的.

证明完毕

状态反馈

如前所述, 若在可控离散时间系统中也采用状态反馈

$$\boldsymbol{u}(k)=\boldsymbol{K}\boldsymbol{x}(k),\ \boldsymbol{K}\in R^{r\times n}, \tag{45}$$

则 $\boldsymbol{F}_k\triangleq(\boldsymbol{F}+\boldsymbol{G}\boldsymbol{K})$ 的特征值可以任意配置. 我们特别感兴趣的是 $\boldsymbol{F}_k$ 的全部特征值均配置成 0 的情况. 这时, 对于某个自然数 $d\leqslant n$, $\boldsymbol{F}_k$ 的最小多项式变成

$$\psi_m(\lambda)=\lambda^d.$$

因此可以看到

$$\boldsymbol{F}_k^d=\mathbf{0},$$

在闭环系统

$$\boldsymbol{x}(k+1)=\boldsymbol{F}_k\boldsymbol{x}(k)$$

中, 对于任意 $\boldsymbol{x}(0)\in R^n$, $\boldsymbol{x}(d)=\mathbf{0}$. 亦即, 在可控离散时间系统中, 适当地选择常数矩阵 $\boldsymbol{K}$, 由任意的起始状态出发, 至多 d 步就可以到达原点. 这是离散时间系统突出的一个特点.

关于观测器[132~136]

与状态反馈相联系, 这里我们简单地讨论一下离散时间系统的观测器.

对于常系数系统

$$\boldsymbol{x}(k+1)=\boldsymbol{F}\boldsymbol{x}(k)+\boldsymbol{G}\boldsymbol{u}(k),\ \boldsymbol{F}\in R^{n\times n},\ \boldsymbol{G}\in R^{n\times r} \tag{46a}$$

$$\boldsymbol{y}(k)=\boldsymbol{C}\boldsymbol{x}(k),\ \boldsymbol{C}\in R^{m\times n}, \tag{46b}$$

同样考虑常系数系统

$$\boldsymbol{z}(k+1)=\widetilde{\boldsymbol{F}}\boldsymbol{z}(k)+\widetilde{\boldsymbol{G}}\boldsymbol{u}(k)+\widetilde{\boldsymbol{H}}\boldsymbol{y}(k),\ \widetilde{\boldsymbol{F}}\in R^{p\times p},\ \widetilde{\boldsymbol{G}}\in R^{p\times r},\ \widetilde{\boldsymbol{H}}\in R^{p\times m}, \tag{47a}$$

$$\boldsymbol{w}(k)=\boldsymbol{E}\boldsymbol{z}(k)+\boldsymbol{D}\boldsymbol{u}(k),\ \boldsymbol{E}\in R^{r\times p},\ \boldsymbol{D}\in R^{r\times r}, \tag{47b}$$

与 $\boldsymbol{x}(0)$, $\boldsymbol{z}(0)$ 值以及输入 $\boldsymbol{u}(\cdot)$ 无关, 当存在某个有限时刻 k_0,

$$\boldsymbol{w}(k)=\boldsymbol{K}\boldsymbol{x}(k),\ \forall k\geqslant k_0 \tag{48}$$

始终成立时, 称系统(47)为 $\boldsymbol{K}\boldsymbol{x}$-观测器. 而且特别是在

$$\boldsymbol{w}(k)=\boldsymbol{x}(k),\ \forall k\geqslant k_0 \tag{49}$$

的情况下, 称为状态观测器. 无论是哪一种情况, 用有限步就可以使 $\boldsymbol{w}$ 的值与 $\boldsymbol{K}\boldsymbol{x}$ 或 $\boldsymbol{x}$ 的值一致, 这是离散时间系统观测器的一个突出特点.

至于状态观测器的构成，和连续时间系统的情况完全一样(参看 A-I-14)，若$(\boldsymbol{F}, \boldsymbol{C})$可观测，则可构成$n$维或$(n-m)$维状态观测器. 例如，根据 A-I-14 中(11)式，可以构成下列

$$\boldsymbol{z}(k+1)=(\boldsymbol{F}-\tilde{\boldsymbol{G}}\boldsymbol{C})\boldsymbol{z}(k)+\tilde{\boldsymbol{G}}\boldsymbol{y}(k)+\boldsymbol{G}\boldsymbol{u}(k) \tag{50a}$$

$$\boldsymbol{w}(k)=\boldsymbol{z}(k) \tag{50b}$$

的n维状态观测器. 对于这种情况下的误差$\tilde{\boldsymbol{z}}\triangleq\boldsymbol{w}-\boldsymbol{x}$，和连续时间系统的情况一样，容易证明

$$\tilde{\boldsymbol{z}}(k+1)=(\boldsymbol{F}-\tilde{\boldsymbol{G}}\boldsymbol{C})\tilde{\boldsymbol{z}}(k) \tag{51}$$

成立. 因此，若这样选择$\tilde{\boldsymbol{G}}$，使得$(\boldsymbol{F}-\tilde{\boldsymbol{G}}\boldsymbol{C})$的特征值均为 0，则对于某个自然数$k_0\leqslant n$，$(\boldsymbol{F}-\tilde{\boldsymbol{G}}\boldsymbol{C})^{k_0}=\boldsymbol{0}$，所以满足$\boldsymbol{w}(k)=\boldsymbol{x}(k)$，${}^\forall k\geqslant k_0$，${}^\forall\boldsymbol{z}(0)\in R^n$，$\boldsymbol{x}(0)\in R^n$，${}^\forall\boldsymbol{u}(\cdot)$的状态观测器的性质.

$(n-m)$维状态观测器也可以和连续时间系统的情况完全一样地构成. 在这种情况下，从起始时刻最多经过$(n-m)$步，观测器的输出$\boldsymbol{w}$便和控制对象的状态$\boldsymbol{x}$一致.

从在尽可能短的时间内产生状态$\boldsymbol{x}$的观点来看，一般可以说$(n-m)$维状态观测器比n维状态观测器更好一些，但是尽管如此，也不能保证最短的时间. 关于这一点，已经知道有下列结果. 即，当$(\boldsymbol{F}, \boldsymbol{C})$可观测时，满足

$$\operatorname{rank}\begin{bmatrix}\boldsymbol{C}\\ \boldsymbol{F}\boldsymbol{C}\\ \vdots\\ \boldsymbol{F}^{\nu-1}\boldsymbol{C}\end{bmatrix}=n$$

的最小整数ν称为可观测性指标(A-I-11). 设$\nu=\beta$为可观测性指标，则可以构成最多$(\beta-1)$步的状态观测器[135].

因为当$\operatorname{rank}\boldsymbol{C}=m$时$\beta-1\leqslant n-m$，则$(n-m)$维状态观测器一般不是最短时间观测器.

显然，若使用上述的$(\beta-1)$步状态观测器，对于任意$\boldsymbol{K}$，则可构成最多经过$(\beta-1)$步就能产生出$\boldsymbol{K}\boldsymbol{x}$值的$\boldsymbol{K}z$-观测器. 当然，对于某个固定的$\boldsymbol{K}$，应该有可能进一步缩短观测步数.

根据最近的研究[132, 133, 136]，关于最短步数问题显然有下列一般结果. 首先，对于给定的$\boldsymbol{K}$，存在以某个有限步数产生$\boldsymbol{K}\boldsymbol{x}$值的离散时间观测器的充分必要条件是1)

$$R[(\boldsymbol{F}^T)^n\boldsymbol{K}^T]\subseteq R[\boldsymbol{C}^T, \boldsymbol{F}^T\boldsymbol{C}^T, \cdots, (\boldsymbol{F}^T)^n\boldsymbol{C}^T]. \tag{52}$$

该条件的成立与$\boldsymbol{F}$的正则性无关，这里我们感兴趣的是$(\boldsymbol{F}, \boldsymbol{C})$的可观测性并不一定必要(即使$\boldsymbol{K}$是正则的). 例如，当$\boldsymbol{F}$是幂零矩阵，$\boldsymbol{F}^\alpha=\boldsymbol{0}$，$0<\alpha\leqslant n$时，上面条件的成立与$(\boldsymbol{F}, \boldsymbol{C})$的可观测性无关.

此外已经知道，在(52)式条件成立的情况下，设满足

$$R[(\boldsymbol{F}^T)^\sigma\boldsymbol{K}^T]\subseteq R[\boldsymbol{C}^T, \boldsymbol{F}^T\boldsymbol{C}^T, \cdots, (\boldsymbol{F}^T)^\sigma\boldsymbol{C}^T] \tag{53}$$

的最小自然数为σ，则σ是$\boldsymbol{K}\boldsymbol{x}$-观测器的最小步数.

也就是说，对于任意起始条件$\boldsymbol{x}(0)$，$\boldsymbol{z}(0)$，过去σ步的输入输出数据$\{\boldsymbol{u}(k), \boldsymbol{y}(k)\}_0^\sigma$对于决定$\boldsymbol{K}\boldsymbol{x}$值是充分必要的. 关于用最短步数产生$\boldsymbol{K}\boldsymbol{x}$的观测器，有若干构成方案，

1) $R(\boldsymbol{A})\subseteq R(\boldsymbol{B})$的意思是，矩阵$\boldsymbol{B}$的值域包含$\boldsymbol{A}$的值域.

详见文献[132～134, 136].

此外，关于使观测器的维数尽可能低的所谓最小维数问题，目前尚未完全解决.

作为特殊情况，若 $\boldsymbol{F}$ 是正则的，则最短步数的条件可以进一步简化成

$$R[\boldsymbol{K}^T]\subseteq R[\boldsymbol{C}^T,\ (\boldsymbol{F}^T)^{-1}\boldsymbol{C}^T,\ \cdots,\ (\boldsymbol{F}^T)^{-\sigma}\boldsymbol{C}^T] \tag{54}$$

而且若 $(\boldsymbol{F},\ \boldsymbol{C})$ 可观测，对于可观测性指标 β，因

$$R[\boldsymbol{C}^T,\ \boldsymbol{F}^T\boldsymbol{C}^T,\ \cdots,\ (\boldsymbol{F}^T)^{\beta-1}\boldsymbol{C}^T]=R^n$$

成立，则对于任何 $\boldsymbol{K}$ 最多经过 $(\beta-1)$ 步就可以决定 $\boldsymbol{Kx}$ 值. 这不外乎是前面讲过的结果.

附录 关于可控性等的引理

[引理] $(\boldsymbol{A},\ \boldsymbol{B})$ 可控

$\Leftrightarrow$设 $\lambda\in C$, $\boldsymbol{\eta}\in C^n$，则除了零向量以外不存在满足 $\boldsymbol{\eta}^T\boldsymbol{A}=\lambda\boldsymbol{\eta}^T$, $\boldsymbol{\eta}^T\boldsymbol{B}=\boldsymbol{0}$ 的 $\boldsymbol{\eta}$.

(证明) $(\Rightarrow)$：设满足上述条件的 $\boldsymbol{\eta}\neq\boldsymbol{0}$ 存在，则

$$\begin{aligned}\boldsymbol{\eta}^T[\boldsymbol{B},\ \boldsymbol{AB},\ \cdots,\ \boldsymbol{A}^{n-1}\boldsymbol{B}]&=[\boldsymbol{\eta}^T\boldsymbol{B},\ \boldsymbol{\eta}^T\boldsymbol{AB},\ \cdots,\ \boldsymbol{\eta}^T\boldsymbol{A}^{n-1}\boldsymbol{B}]\\&=[\boldsymbol{\eta}^T\boldsymbol{B},\ \lambda\boldsymbol{\eta}^T\boldsymbol{B},\ \cdots,\ \lambda^{n-1}\boldsymbol{\eta}^T\boldsymbol{B}]=\boldsymbol{0}\end{aligned}$$

所以

$$\mathrm{Re}\{\boldsymbol{\eta}^T[\boldsymbol{B},\ \boldsymbol{AB},\ \cdots,\ \boldsymbol{A}^{n-1}\boldsymbol{B}]\}=\boldsymbol{0}.$$

因矩阵 $\boldsymbol{A}$, $\boldsymbol{B}$ 都是实矩阵，由上式则

$$\{\mathrm{Re}\,\boldsymbol{\eta}^T\}[\boldsymbol{B},\ \boldsymbol{AB},\ \cdots,\ \boldsymbol{A}^{n-1}\boldsymbol{B}]=\boldsymbol{0}$$

成立. 即矩阵 $[\boldsymbol{B},\ \boldsymbol{AB},\cdots,\ \boldsymbol{A}^{n-1}\boldsymbol{B}]$ 的行(在实数体上)不是线性独立的. 因此 $(\boldsymbol{A},\ \boldsymbol{B})$ 不是可控的.

$(\Leftarrow)$：设 $\boldsymbol{A}$ 的若当形为 $\boldsymbol{J}=\boldsymbol{T}^{-1}\boldsymbol{AT}$，且 $\widetilde{\boldsymbol{B}}\triangleq\boldsymbol{T}^{-1}\boldsymbol{B}$. 利用 A-I-11 中定理 4, 若假定 $(\boldsymbol{A},\ \boldsymbol{B})$ 不是可控的，对于某个 $j\,(1\leqslant j\leqslant\sigma)$，则

$$\mathrm{rank}\begin{bmatrix}\boldsymbol{\beta}_{j1}\\ \vdots\\ \boldsymbol{\beta}_{j\alpha_j}\end{bmatrix}^{*}<\alpha_j.$$

因此，有满足 $\sum\limits_{k=1}^{\alpha_j}C_{jk}\boldsymbol{\beta}_{jk}=\boldsymbol{0}$ 的不全为 0 的 $C_{jk}\,(k=1,\ \cdots,\ \alpha_j)$ 存在. 设 $\boldsymbol{\beta}_{jk}$ 是 $\widetilde{\boldsymbol{B}}$ 的第 d_k 行 $(k=1,\ \cdots,\ \alpha_j)$，向量 $\hat{\boldsymbol{\eta}}$ 为

$$\hat{\boldsymbol{\eta}}=[0\cdots0\underset{\substack{\uparrow\\ \text{第 } d_1 \text{ 个元素}}}{C_{j1}}\vdots0\cdots0\underset{\substack{\uparrow\\ \text{第 } d_2 \text{ 个元素}}}{C_{j2}}\vdots\cdots\vdots0\cdots0\underset{\substack{\uparrow\\ \text{第 } d_{\alpha_j} \text{ 个元素}}}{C_{j\alpha_j}}\vdots0\cdots0]^T,$$

则 $\hat{\boldsymbol{\eta}}\neq\boldsymbol{0}$, $\hat{\boldsymbol{\eta}}^T\widetilde{\boldsymbol{B}}=0$. 而且根据若当形的构成，$\hat{\boldsymbol{\eta}}^T\boldsymbol{J}=\lambda_j\hat{\boldsymbol{\eta}}^T$ 成立，其中 λ_j 是 $\boldsymbol{A}$ 的特征值. 因此，若令 $\boldsymbol{\eta}\triangleq(\boldsymbol{T}^{-1})^T\hat{\boldsymbol{\eta}}$，则

$$\boldsymbol{\eta}^T\boldsymbol{B}=\hat{\boldsymbol{\eta}}^T\boldsymbol{T}^{-1}\boldsymbol{B}=\hat{\boldsymbol{\eta}}^T\widetilde{\boldsymbol{B}}=\boldsymbol{0}$$

$$\boldsymbol{\eta}^T\boldsymbol{A}=\hat{\boldsymbol{\eta}}^T\boldsymbol{T}^{-1}\boldsymbol{A}=\hat{\boldsymbol{\eta}}^T\boldsymbol{J}\boldsymbol{T}^{-1}=\lambda_j\hat{\boldsymbol{\eta}}^T\boldsymbol{T}^{-1}=\lambda_j\boldsymbol{\eta}^T$$

成立.

证明完毕

---

* 原文误为 $\mathrm{rank}\begin{bmatrix}\boldsymbol{\beta}_{j1}\\ \vdots\\ \boldsymbol{\beta}_{i\alpha_j}\end{bmatrix}<\alpha_j$. ——译者注

作为该引理的对偶定理，显然下列结果成立.

[系 1]　$(\boldsymbol{C}, \boldsymbol{A})$ 可观测

$\Leftrightarrow$设 $\lambda\in C$, $\boldsymbol{\xi}\in C^n$, 则除了零向量之外不存在满足 $\boldsymbol{A\xi}=\lambda\boldsymbol{\xi}$, $\boldsymbol{C\xi}=\boldsymbol{0}$ 的 $\boldsymbol{\xi}$.

再则，显然上述 $\lambda\in C$ 是 $\boldsymbol{A}$ 的特征值. 考虑到这一点可得下列结果.

[系 2]　$(\boldsymbol{A}, \boldsymbol{B})$ 可稳定

$\Leftrightarrow$设 $\lambda\in C$, $\boldsymbol{\eta}\in C^n$, 若有满足 $\boldsymbol{\eta}^T\boldsymbol{A}=\lambda\boldsymbol{\eta}^T$, $\boldsymbol{\eta}^T\boldsymbol{B}=\boldsymbol{0}$ 的不是零向量的 $\boldsymbol{\eta}$ 存在，则 $\mathrm{Re}(\lambda)<0$.

(证明)　根据 P-I-5 中定理 1, 可稳定的充分必要条件是，由 A-I-11 中定理 5 所用过的正则变换变换成

$$\boldsymbol{A}^c=\begin{bmatrix}\boldsymbol{A}_{11}^c & \boldsymbol{A}_{12}^c\\ \boldsymbol{0} & \boldsymbol{A}_{22}^c\end{bmatrix}\begin{matrix}\}q^c\\ \}n-q^c,\end{matrix}\qquad \boldsymbol{B}^c=\begin{bmatrix}\boldsymbol{B}_1^c\\ \boldsymbol{0}\end{bmatrix}\begin{matrix}\}q^c\\ \}n-q^c\end{matrix}$$

$$(\boldsymbol{A}_{11}^c,\ \boldsymbol{B}_1^c)\text{ 可控}$$

时，$\mathrm{Re}\lambda(\boldsymbol{A}_{22}^c)<0$.

$(\Rightarrow)$：假定满足 $\boldsymbol{\eta}^T\boldsymbol{A}^c=\lambda\boldsymbol{\eta}^T$, $\boldsymbol{\eta}^T\boldsymbol{B}^c=\boldsymbol{0}$, $\mathrm{Re}(\lambda)\geqslant 0$ 的 $\boldsymbol{\eta}\neq\boldsymbol{0}$ 存在，将该 $\boldsymbol{\eta}$ 分割成

$$\boldsymbol{\eta}=\begin{bmatrix}\boldsymbol{\eta}_1\\ \boldsymbol{\eta}_2\end{bmatrix}\begin{matrix}\}q^c\\ \}n-q^c,\end{matrix}$$

则

$$\boldsymbol{\eta}^T\boldsymbol{A}^c=[\boldsymbol{\eta}_1^T\boldsymbol{A}_{11}^c,\ \boldsymbol{\eta}_1^T\boldsymbol{A}_{12}^c+\boldsymbol{\eta}_2^T\boldsymbol{A}_{22}^c]=[\lambda\boldsymbol{\eta}_1,\ \lambda\boldsymbol{\eta}_2],\ \boldsymbol{\eta}^T\boldsymbol{B}^c=\boldsymbol{\eta}_1^T\boldsymbol{B}_1^c=\boldsymbol{0}.$$

根据上列两式及 $(\boldsymbol{A}_{11}^c, \boldsymbol{B}_1^c)$ 可控，则 $\boldsymbol{\eta}_1=\boldsymbol{0}$. 将其代入上面第 1 式，得

$$\boldsymbol{\eta}_2^T\boldsymbol{A}_{22}^c=\lambda\boldsymbol{\eta}_2.$$

该式表示 λ 是 $\boldsymbol{A}_{22}^c$ 的特征值，因 $\mathrm{Re}(\lambda)\geqslant 0$, 则不是可稳定的.

$(\Leftarrow)$：若不是可稳定的，则在 $\boldsymbol{A}_{22}^c$ 的特征值中有 $\mathrm{Re}(\lambda)\geqslant 0$ 的特征值存在. 设与此特征值相对应的 $\boldsymbol{A}_{22}^c$ 的特征向量 ($(n-q^c)$ 维向量) 为 $\boldsymbol{\eta}_2$, 若选成

$$\boldsymbol{\eta}=\begin{bmatrix}\boldsymbol{0}\\ \boldsymbol{\eta}_2\end{bmatrix}\begin{matrix}\}q^c\\ \}n-q^c,\end{matrix}$$

显然 $\boldsymbol{\eta}\neq\boldsymbol{0}$, 而且 $\boldsymbol{\eta}^T\boldsymbol{A}^c=\lambda\boldsymbol{\eta}^T$, $\boldsymbol{\eta}^T\boldsymbol{B}^c=\boldsymbol{0}$ 成立.

证明完毕

[系 3]　$(\boldsymbol{C}, \boldsymbol{A})$ 可检测

$\Leftrightarrow$设 $\lambda\in C$, $\boldsymbol{\xi}\in C^n$, 若有满足 $\boldsymbol{A\xi}=\lambda\boldsymbol{\xi}$, $\boldsymbol{C\xi}=\boldsymbol{0}$ 的不是零向量的 $\boldsymbol{\xi}$ 存在，则 $\mathrm{Re}(\lambda)<0$.

B-I-15 矩阵方程式

本章将讲述在包含有未知矩阵的方程式中对于自动控制中的应用特别重要的一些方程式. 因以在自动控制中的应用为前提,故假定矩阵的元素均为实数.

其1 线性方程式

1.1 微分方程式

首先考虑线性矩阵微分方程式

$$\dot{\boldsymbol{X}}=\boldsymbol{X}\boldsymbol{D}(t)+\boldsymbol{E}(t)\boldsymbol{X}+\boldsymbol{F}(t) \quad \boldsymbol{X}(t_0)=\boldsymbol{X}_0, \tag{1}$$

其中假定, $\boldsymbol{X}=m\times n$ 为未知矩阵, $\boldsymbol{D}(\cdot)$, $\boldsymbol{E}(\cdot)$, $\boldsymbol{F}(\cdot)$ 分别为适当阶数的矩阵,其元素是 t 的分段连续函数.

[引理 1] 对于任意的 $t\in R$, (1)式的唯一解 $\boldsymbol{X}(t;t_0,\boldsymbol{X}_0)$ 存在,且可以用

$$\boldsymbol{X}(t;t_0,\boldsymbol{X}_0)=\boldsymbol{\Phi}_1(t,t_0)\boldsymbol{X}_0\boldsymbol{\Phi}_2(t,t_0)+\int_{t_0}^{t}\boldsymbol{\Phi}_1(t,\tau)\boldsymbol{F}(\tau)\boldsymbol{\Phi}_2(t,\tau)d\tau$$

给出. 其中 $\boldsymbol{\Phi}_1(t,t_0)$, $\boldsymbol{\Phi}_2(t,t_0)$ 是分别用

$$\frac{\partial\boldsymbol{\Phi}_1}{\partial t}=\boldsymbol{E}(t)\boldsymbol{\Phi}_1 \quad \boldsymbol{\Phi}_1(t_0,t_0)=\boldsymbol{I}_m$$

$$\frac{\partial\boldsymbol{\Phi}_2}{\partial t}=\boldsymbol{\Phi}_2\boldsymbol{D}(t) \quad \boldsymbol{\Phi}_2(t_0,t_0)=\boldsymbol{I}_n$$

定义的状态转移矩阵(参考第 203 页 A-I-6 中问题 6).

在该引理中特别是若限定

$$\begin{aligned}&n=m\\&\boldsymbol{D}(t)=\boldsymbol{A}(\text{常数矩阵})\\&\boldsymbol{E}(t)=\boldsymbol{A}^T\\&\boldsymbol{F}(t)=\boldsymbol{Q}(\text{常数对称矩阵})\end{aligned}$$

则可得下列系.

[系 1] 对于任意 $t\in R$, 线性常系数矩阵微分方程式

$$\dot{\boldsymbol{X}}=\boldsymbol{X}\boldsymbol{A}+\boldsymbol{A}^T\boldsymbol{X}+\boldsymbol{Q} \quad \boldsymbol{X}(t_0)=\boldsymbol{X}_0 \tag{2}$$

的唯一解 $\boldsymbol{X}(t;t_0,\boldsymbol{X}_0)$ 存在,可用

$$\boldsymbol{X}(t;t_0,\boldsymbol{X}_0)=e^{\boldsymbol{A}^T(t-t_0)}\boldsymbol{X}_0e^{\boldsymbol{A}(t-t_0)}+\int_{t_0}^{t}e^{\boldsymbol{A}^T(t-\tau)}\boldsymbol{Q}e^{\boldsymbol{A}(t-\tau)}d\tau \tag{3}$$

给出,且满足下列性质

$$\boldsymbol{X}(t;t_0,\boldsymbol{X}_0)=[\boldsymbol{X}(t;t_0,\boldsymbol{X}_0)]^T\in R^{n\times n} \quad \forall\boldsymbol{X}_0=\boldsymbol{X}_0^T\in R^{n\times n}. \tag{4}$$

[系 2] 若 $\boldsymbol{A}$ 渐近稳定(即 $\mathrm{Re}\,\lambda_i(\boldsymbol{A})<0$, $i=1, 2, \cdots, n$)[1], 则对于任意实对称矩阵 $\boldsymbol{X}_0\in R^{n\times n}$, (2)式解 $\boldsymbol{X}(t;t_0,\boldsymbol{X}_0)$ 在 $t\geqslant t_0$ 时一致有界[2].

1) 以后常简单地将其表示成 $\mathrm{Re}\lambda(\boldsymbol{A})<0$.

2) 所谓实对称矩阵 $\boldsymbol{X}(t)$ 对于 $t_0\leqslant t\leqslant t_1$ 一致上(下)有界, 就是存在与 t 无关的某个实对称矩阵 $\boldsymbol{K}_1\in R^{n\times n}$($\boldsymbol{K}_2\in R^{n\times n}$),使得 $\boldsymbol{X}(t)\leqslant\boldsymbol{K}_1$($\boldsymbol{X}(t)\geqslant\boldsymbol{K}_2$) $t_0\leqslant t\leqslant t_1$ 成立. 而且当上、下都一致有界时,只称为一致有界.

（证明） 由于解用(3)式给定，及矩阵范数的性质（参看第159～161页），则

$$\|X(t;t_0,X_0)\|\leqslant\|e^{A^T(t-t_0)}X_0e^{A(t-t_0)}\|+\left\|\int_{t_0}^{t}e^{A^T(t-t_0)}Qe^{A(t-t_0)}dt\right\|$$

$$\leqslant\|X_0\|\|e^{A(t-t_0)}\|^2+\int_{t_0}^{t}\|Q\|\|e^{A(t-t_0)}\|^2dt\quad {}^\forall X_0=X_0^T\in R^{n\times n}\quad {}^\forall t\in R$$

成立[1]. 因其中 $\mathrm{Re}\lambda(A)<0$, 则存在某个常数 $a>0$ 和 $\lambda>0$,

$$\|e^{A(t-t_0)}\|<ae^{-\lambda(t-t_0)}\leqslant a\quad {}^\forall t\geqslant t_0$$

成立（参看P-I-5中(11)式）. 因此

$$\|X(t;t_0,X_0)\|\leqslant a^2\|X_0\|+a^2\|Q\|\int_0^\infty e^{-2\lambda\tau}d\tau$$

$$=a^2\left(\|X_0\|+\frac{1}{2\lambda}\|Q\|\right)\triangleq\alpha(X_0)\quad {}^\forall X_0=X_0^T\in R^{n\times n}\quad {}^\forall t\geqslant t_0$$

而且，由(4)式，因 $X(t;t_0,X_0)$ 是实对称矩阵，根据矩阵范数的性质（参看A-I-11中附录3的(ii)(b)[2]），则

$$-\alpha(X_0)I_n\leqslant X(t;t_0,X_0)\leqslant\alpha(X_0)I_n\quad {}^\forall t\geqslant t_0$$

成立. 因为当 X_0 固定时 $\alpha(X_0)$ 有界，而且是与 t 无关的常数，故得到该系.

证明完毕

1.2 代数方程式

现在我们来讨论对于给定的矩阵 $D\in R^{n\times n}$, $E\in R^{m\times m}$, $F\in R^{m\times n}$ 求满足线性矩阵方程式

$$XD+EX=-F\tag{5}$$

的矩阵 $X\in R^{m\times n}$ 的问题.

首先，关于(5)式解 X 的唯一性有下列性质.

［引理2］ (5)式解唯一确定的充分必要条件是，D 及 E 的特征值满足

$$\lambda_i(D)+\lambda_j(E)\neq0\quad i=1,2,\cdots,n;\ j=1,2,\cdots,m\tag{6}$$

（证明） 将(5)式两边列展开，可见(5)式和下列线性方程组

$$(D^T\otimes I_m+I_n\otimes E)\mathrm{CS}\,X=-\mathrm{CS}\,F\tag{7}$$

等价（参照B-I-4）.

我们知道（参照B-I-11），关于矩阵 $(D^T\otimes I_m+I_nE)\in R^{nm\times nm}$ 的特征值，下式

$$\left.\begin{aligned}&\lambda_k(D^T\otimes I_m+I_n\otimes E)=\lambda_i(D)+\lambda_j(E)\\&k=1,2,\cdots,nm;\ i=1,2,\cdots,n;\ j=1,2,\cdots,m\end{aligned}\right\}\tag{8}$$

成立.

显然，(8)式左边不为零，即(6)式成立（参照B-I-4及B-I-11中系3），是线性方程组(7)的解唯一存在的充分必要条件.

证明完毕

1) 利用对于欧几里得范数$\|\cdot\|$, $\|A\|=\|A^T\|$ 成立的性质（参看A-I-11中的附录）.

2) 在A-I-11中附录3的注意(ii)(b)已证明，对于任意实对称矩阵 S 和正实数 α, $\|S\|<\alpha\Rightarrow S\leqslant\alpha I$ 成立. 在 $\|\cdot\|$ 是由向量范数导出的矩阵范数的情况下，根据B-I-11中定理13，可以用和上述注意中同样的方法证明，$\|S\|\leqslant\alpha\Rightarrow-\alpha I\leqslant S\leqslant\alpha I$ 成立.

在这里，特别是若设 $m=n$, $\boldsymbol{D}=\boldsymbol{A}$, $\boldsymbol{E}=\boldsymbol{A}^T$, $\boldsymbol{F}=\boldsymbol{Q}$(其中 $\boldsymbol{Q}=\boldsymbol{Q}^T$)，则(5)式变成

$$\boldsymbol{XA}+\boldsymbol{A}^T\boldsymbol{X}=-\boldsymbol{Q}. \tag{9}$$

因为这种形式的线性矩阵方程式在稳定性理论等方面的应用上特别重要，所以下面我们来讨论一下(9)式解的求法及其性质.

首先，根据 $\boldsymbol{Q}$ 的对称性容易证明，若(9)式的解存在，则必存在实对称解(若 $\boldsymbol{X}\in R^{n\times n}$ 是解，则 $\frac{1}{2}(\boldsymbol{X}+\boldsymbol{X}^T)$ 也是解). 因此，若(9)式具有唯一解 $\boldsymbol{X}\in R^{n\times n}$，则其为实对称解. 于是关于(9)式的唯一解有下列定理.

[定理 1] 若 $\boldsymbol{A}$ 渐近稳定，则(9)式具有唯一解 $\boldsymbol{X}$，关于 $\boldsymbol{X}$ 有下列性质.

(i) $\boldsymbol{X}$ 是可以用

$$\boldsymbol{X}=\int_0^\infty e^{\boldsymbol{A}^Tt}\boldsymbol{Q}e^{\boldsymbol{A}t}\,dt \tag{10}$$

给出的有界实对称矩阵.

(ii) $\boldsymbol{Q}>\boldsymbol{0}(<\boldsymbol{0})\Rightarrow\boldsymbol{X}>\boldsymbol{0}(<\boldsymbol{0})$

(iii) $\boldsymbol{Q}\geqslant\boldsymbol{0}(\leqslant\boldsymbol{0})\Rightarrow\boldsymbol{X}\geqslant\boldsymbol{0}(\leqslant\boldsymbol{0})$

(iv) 在(5)式中设 $\boldsymbol{Q}=\boldsymbol{Q}_1(\boldsymbol{Q}=\boldsymbol{Q}_2)$，令这时的解为 $\boldsymbol{X}_1(\boldsymbol{X}_2)$，则

$$\boldsymbol{Q}_1>\boldsymbol{Q}_2\Rightarrow\boldsymbol{X}_1>\boldsymbol{X}_2$$

(v) $\boldsymbol{Q}_1\geqslant\boldsymbol{Q}_2\Rightarrow\boldsymbol{X}_1\geqslant\boldsymbol{X}_2$

(vi) 若 $\boldsymbol{Q}\geqslant\boldsymbol{0}$，则存在适当的矩阵 $\boldsymbol{C}\in R^{r\times n}$($r\triangleq\operatorname{rank}\boldsymbol{Q}$)，可以表示成 $\boldsymbol{Q}=\boldsymbol{C}^T\boldsymbol{C}$. 这时

$$\text{矩阵对}(\boldsymbol{C},\ \boldsymbol{A})\text{可观测}\Rightarrow\boldsymbol{X}>\boldsymbol{0}$$

(证明) 首先，根据引理 2 显然具有唯一解.

(i) 根据系 1, 系 2, 因为对于 $t\geqslant0$, $\boldsymbol{X}(t;0,\boldsymbol{0})=\int_0^t e^{\boldsymbol{A}^Tt}\boldsymbol{Q}e^{\boldsymbol{A}t}\,dt$ 一致有界，显然(10)式右边是有界的. 其次，将被积分项对 t 微分(参照 B-I-13 中定理 2, 例 6)，得

$$\frac{d}{dt}\{e^{\boldsymbol{A}^Tt}\boldsymbol{Q}e^{\boldsymbol{A}t}\}=\boldsymbol{A}^Te^{\boldsymbol{A}^Tt}\boldsymbol{Q}e^{\boldsymbol{A}t}+e^{\boldsymbol{A}^Tt}\boldsymbol{Q}e^{\boldsymbol{A}t}\boldsymbol{A},$$

将该式两边从 0 到 ∞ 积分，得

$$\boldsymbol{A}^T\left(\int_0^\infty e^{\boldsymbol{A}^Tt}\boldsymbol{Q}e^{\boldsymbol{A}t}\,dt\right)+\left(\int_0^\infty e^{\boldsymbol{A}^Tt}\boldsymbol{Q}e^{\boldsymbol{A}t}\,dt\right)\boldsymbol{A}=\left[e^{\boldsymbol{A}^Tt}\boldsymbol{Q}e^{\boldsymbol{A}t}\right]_0^\infty=-\boldsymbol{Q},$$

即 $\boldsymbol{X}$ 可由(10)式给出，因此是实对称的.

(ii) 设 $\boldsymbol{Q}>\boldsymbol{0}(<\boldsymbol{0})$，在(10)式两边右乘以向量 $\boldsymbol{y}\in R^n$，左乘以 $\boldsymbol{y}^T$，得

$$\begin{aligned}\boldsymbol{y}^T\boldsymbol{X}\boldsymbol{y}&=\int_0^\infty(e^{\boldsymbol{A}t}\boldsymbol{y})^T\boldsymbol{Q}(e^{\boldsymbol{A}t}\boldsymbol{y})\,dt\\&>0(<0)\qquad \boldsymbol{y}\neq\boldsymbol{0}\\&=0\qquad\qquad\ \ \boldsymbol{y}=\boldsymbol{0}\\&\therefore\quad \boldsymbol{X}>\boldsymbol{0}(<\boldsymbol{0})\end{aligned}$$

(iii) 和(ii)一样.

(iv), (v) 首先由(9)式，$(\boldsymbol{X}_1-\boldsymbol{X}_2)\boldsymbol{A}+\boldsymbol{A}^T(\boldsymbol{X}_1-\boldsymbol{X}_2)=-(\boldsymbol{Q}_1-\boldsymbol{Q}_2)$ 成立. 由此及(ii), (iii)显然.

(vi) 因实对称矩阵与对角矩阵正交相似，则可以表示成 $\boldsymbol{Q}=\boldsymbol{C}^T\boldsymbol{C}$(参看 B-I-11 中系 20).

而且我们知道，若矩阵对 $(\boldsymbol{C},\ \boldsymbol{A})$ 可观测，则对于某个 $t_f>0$，准正定对称矩阵

$$\boldsymbol{W}_0(0,\ t_f)\triangleq\int_0^{t_f}e^{\boldsymbol{A}^Tt}\boldsymbol{C}^T\boldsymbol{C}e^{\boldsymbol{A}t}\,dt \tag{11}$$

为正则的，即

$$\boldsymbol{W}_0(0,\ t_f)>\boldsymbol{0} \tag{12}$$

成立(参看 A-I-8 中定理 2).

由(i)及(11)，(12)式，得

$$\boldsymbol{X}\geqslant\boldsymbol{W}_0(0,\ t_f)>\boldsymbol{0}.$$

证明完毕

(注意)

“(a) 性质(i)不限于 $n=m$，$\boldsymbol{D}=\boldsymbol{A}$，$\boldsymbol{E}=\boldsymbol{A}^T$，$\boldsymbol{F}=\boldsymbol{Q}$，对于更一般的情况也成立. 即，若 $\boldsymbol{D}$，$\boldsymbol{E}$ 均渐近稳定，则(5)式的唯一解 $\boldsymbol{X}$ 可以用

$$\boldsymbol{X}=\int_0^{\infty}e^{\boldsymbol{E}t}\boldsymbol{F}e^{\boldsymbol{D}t}\,dt$$

给出(试证明).

(b) 根据定理 1，若 $\boldsymbol{A}$ 渐近稳定，则性质(i)～(vi)成立. 其逆，即假定方程式(9)的解已求出，若它具有(i)～(vi)的性质，$\mathrm{Re}\lambda(\boldsymbol{A})<0$ 是否真的成立. 关于这个问题我们再稍微讲一下.”

(ii)，(iv)的逆照样成立.

(vii) “若 $\boldsymbol{Q}>\boldsymbol{0}(<\boldsymbol{0})$ 且解 $\boldsymbol{X}$ 为对称正定(负定)，则 $\boldsymbol{A}$ 是渐近稳定的.”

(viii) “在(9)式中设 $\boldsymbol{Q}=\boldsymbol{Q}_1(\boldsymbol{Q}=\boldsymbol{Q}_2)$ 时的解为 $\boldsymbol{X}_1(\boldsymbol{X}_2)$，若 $\boldsymbol{Q}_1>\boldsymbol{Q}_2$ 且 $\boldsymbol{X}_1>\boldsymbol{X}_2$，则 $\boldsymbol{A}$ 渐近稳定.”

关于(vi)，比其逆的条件稍强一些的如下性质成立.

(ix) “若可以表示成 $\boldsymbol{Q}=\boldsymbol{C}^T\boldsymbol{C}$，$(\boldsymbol{C},\ \boldsymbol{A})$ 可检测(参看 A-I-16 后面附录中系 3)，而且解 $\boldsymbol{X}$ 为对称准正定(准负定)，则 $\mathrm{Re}\lambda(\boldsymbol{A})<0(>0)$.”

(证明)

(vii) 设 $\boldsymbol{A}$ 具有满足 $\mathrm{Re}\,\lambda\geqslant0$ 的特征值 $\lambda(\mathrm{Re}\,\lambda\geqslant0)$，对应于 λ 的 $\boldsymbol{A}$ 的特征向量为 $\boldsymbol{x}\in C^n(\boldsymbol{x}\neq\boldsymbol{0})$，则 $\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$，$\boldsymbol{x}^*\boldsymbol{A}^T=\bar{\lambda}\boldsymbol{x}^*$. 因此，将(9)式左乘以 $\boldsymbol{x}^*$，右乘以 $\boldsymbol{x}$，整理后得

$$2\,\mathrm{Re}\,\lambda\boldsymbol{x}^*\boldsymbol{X}\boldsymbol{x}=-\boldsymbol{x}^*\boldsymbol{Q}\boldsymbol{x}, \tag{13}$$

因其中假定 $\mathrm{Re}\,\lambda\geqslant0$ 且 $\boldsymbol{X}$，$\boldsymbol{Q}$ 均为对称正定(负定)，则该式左边非负(非正)，右边为负(正)，互相矛盾.

(viii) 设 $\varDelta\boldsymbol{X}\triangleq\boldsymbol{X}_1-\boldsymbol{X}_2$，$\varDelta\boldsymbol{Q}\triangleq\boldsymbol{Q}_1-\boldsymbol{Q}_2$，由(9)式，则

$$\varDelta\boldsymbol{X}\boldsymbol{A}+\boldsymbol{A}^T\varDelta\boldsymbol{X}=-\varDelta\boldsymbol{Q} \tag{14}$$

成立. 因根据假定 $\varDelta\boldsymbol{X}$，$\varDelta\boldsymbol{Q}$ 均为对称正定的，由(vii)，则 $\boldsymbol{A}$ 渐近稳定.

(ix) 设 $\boldsymbol{A}$ 具有满足 $\mathrm{Re}\,\lambda\geqslant0(\leqslant0)$ 的特征值 λ，则和(vii)完全一样，可得(13)式. 因此，若令 $\boldsymbol{Q}=\boldsymbol{C}^T\boldsymbol{C}$，则

$$2\mathrm{Re}\,\lambda\boldsymbol{x}^*\boldsymbol{X}\boldsymbol{x}+\boldsymbol{x}^*\boldsymbol{C}^T\boldsymbol{C}\boldsymbol{x}=0$$

成立. 因其中假定 $\mathrm{Re}\,\lambda\geqslant 0(\leqslant 0)$, $\boldsymbol{X}\geqslant\boldsymbol{0}(\leqslant\boldsymbol{0})$, 则左边第1, 第2项都是非负的. 因此, 为使该式成立必须

$$\boldsymbol{Cx}=\boldsymbol{0}.$$

这与可检测性相矛盾(参看 A-I-16 中附录系 3).

证明完毕

性质(vi)的逆本来不成立, 但加上下列条件后则成立.

(x) "若矩阵对 $(\boldsymbol{C}, \boldsymbol{A})$ 可检测, 而且

$$\boldsymbol{X}\triangleq\int_0^\infty e^{\boldsymbol{A}^T t}\boldsymbol{C}^T\boldsymbol{C}e^{\boldsymbol{A}t}\,dt \tag{15}$$

有界, 则 $\boldsymbol{A}$ 渐近稳定."

(证明) 设 $\boldsymbol{A}$ 具有满足 $\mathrm{Re}\,\lambda\geqslant 0$ 的特征值 λ, 对应于 λ 的 $\boldsymbol{A}$ 的特征向量为 $\boldsymbol{x}(\boldsymbol{x}\neq\boldsymbol{0})$, 则 $\boldsymbol{Ax}=\lambda\boldsymbol{x}$. 在(15)式右乘以 $\boldsymbol{x}$, 左乘以 $\boldsymbol{x}^*$, 则

$$\int_0^\infty e^{2t\,\mathrm{Re}\,\lambda}\|\boldsymbol{Cx}\|^2 dt<\infty.$$

因其中根据假定 $\mathrm{Re}\,\lambda\geqslant 0$, 则为使该式成立必须

$$\boldsymbol{Cx}=\boldsymbol{0}.$$

这与可检测性相矛盾.

证明完毕

性质(iii)和(v)的逆也不成立, 但是下列两个性质成立.

(xi) "若 $\boldsymbol{Q}\geqslant\boldsymbol{0}$ 且解 $\boldsymbol{X}$ 对称正定, 则 $\boldsymbol{A}$ 稳定(即 $\mathrm{Re}\,\lambda_i(\boldsymbol{A})\leqslant 0$ $i=1, 2, \cdots, n$."

(xii) "在(9)式中设 $\boldsymbol{Q}=\boldsymbol{Q}_1(\boldsymbol{Q}=\boldsymbol{Q}_2)$ 时的解为 $\boldsymbol{X}_1(\boldsymbol{X}_2)$, 若 $\boldsymbol{Q}_1\geqslant\boldsymbol{Q}_2$ 且 $\boldsymbol{X}_1>\boldsymbol{X}_2$, 则 $\boldsymbol{A}$ 稳定."

在这里要注意, 代替 $\boldsymbol{X}\geqslant\boldsymbol{0}(\boldsymbol{X}_1\geqslant\boldsymbol{X}_2)$, 假定比它更强的条件 $\boldsymbol{X}>\boldsymbol{0}(\boldsymbol{X}_1>\boldsymbol{X}_2)$, 也不能保证 $\mathrm{Re}\,\lambda_i(\boldsymbol{A})<0$, 仅限于证明比它弱的性质 $\mathrm{Re}\,\lambda_i(\boldsymbol{A})\leqslant 0$.

(证明)

(xi) 在性质(vii)的证明中将 $\mathrm{Re}\,\lambda\geqslant 0$ 变成 $\mathrm{Re}\,\lambda>0$, 将 $\boldsymbol{Q}>\boldsymbol{0}(\boldsymbol{Q}<\boldsymbol{0})$ 变成 $\boldsymbol{Q}\geqslant\boldsymbol{0}(\boldsymbol{Q}\leqslant\boldsymbol{0})$, 则同样可以证明是矛盾的.

(xii) 由(14)式和(xi)显然.

证明完毕

1.3 平衡点和稳定性

若代数方程式(9)的解存在, 显然它是微分方程式(2)的平衡点. 因此, 下面来考虑该平衡点的稳定性.

首先, 容易证明下列性质.

[引理 3] 对于任意 $\boldsymbol{X}_0\in R^{n\times n}$, 微分方程式(2)的定常解 $\lim\limits_{t\to\infty}\boldsymbol{X}(t; 0, \boldsymbol{X}_0)\triangleq\overline{\boldsymbol{X}}$ 存在的充分必要条件是, $\mathrm{Re}\,\lambda(\boldsymbol{A})<0$(试证明).

由该性质和微分方程式(2)的定常解 $\overline{\boldsymbol{X}}$ 是代数方程式(9)的解(因而也是(2)式的平衡点), 以及定理 1 可见, 这时平衡点 $\overline{\boldsymbol{X}}$ 具有下列性质.

(i) $\overline{\boldsymbol{X}}$ 是大范围渐近稳定的.

(ii) $\bar{\boldsymbol{X}}$ 是代数方程式(9)的唯一解.

(iii) $\boldsymbol{Q}>0(\boldsymbol{Q}<0)\Rightarrow\bar{\boldsymbol{X}}>0(\bar{\boldsymbol{X}}<0)$.

(iv) $\boldsymbol{Q}\geqslant0(\boldsymbol{Q}\leqslant0)\Rightarrow\bar{\boldsymbol{X}}\geqslant0(\bar{\boldsymbol{X}}\leqslant0)$.

(v) $\boldsymbol{Q}_1>\boldsymbol{Q}_2\Rightarrow\bar{\boldsymbol{X}}_1>\bar{\boldsymbol{X}}_2$.

(vi) $\boldsymbol{Q}_1\geqslant\boldsymbol{Q}_2\Rightarrow\bar{\boldsymbol{X}}_1\geqslant\bar{\boldsymbol{X}}_2$.

(vii) $\boldsymbol{Q}=\boldsymbol{C}^T\boldsymbol{C}$, $(\boldsymbol{C}, \boldsymbol{A})$可观测$\Rightarrow\bar{\boldsymbol{X}}>0$.

1.4 线性方程式 $-\boldsymbol{X}+\boldsymbol{EXD}=-\boldsymbol{F}$

下面我们来讨论，对于给定矩阵 $\boldsymbol{D}\in R^{n\times n}$, $\boldsymbol{E}\in R^{m\times m}$, $\boldsymbol{F}\in R^{m\times n}$，满足线性方程式

$$-\boldsymbol{X}+\boldsymbol{EXD}=-\boldsymbol{F} \tag{16}$$

的矩阵 $\boldsymbol{X}\in R^{m\times n}$ 的性质.

首先，关于(16)式解的唯一性有下列性质.

[引理 4] (16)式解 $\boldsymbol{X}$ 唯一确定的充分必要条件是，$\boldsymbol{D}$ 及 $\boldsymbol{E}$ 的特征值满足下列关系

$$\lambda_i(\boldsymbol{D})\lambda_j(\boldsymbol{E})\neq1 \quad i=1,2,\cdots,n; j=1,2,\cdots,m.$$

(证明) 和引理 1 的证明一样. 首先，因(16)式和

$$(-\boldsymbol{I}_n\otimes\boldsymbol{I}_m+\boldsymbol{D}^T\otimes\boldsymbol{E})\operatorname{cs}\boldsymbol{X}=-\operatorname{cs}\boldsymbol{F}$$

等价，而且$(-\boldsymbol{I}_n\otimes\boldsymbol{I}_m+\boldsymbol{D}^T\otimes\boldsymbol{E})$的特征值可以表示成

$$\lambda_k(-\boldsymbol{I}_n\otimes\boldsymbol{I}_m+\boldsymbol{D}^T\otimes\boldsymbol{E})=-1+\lambda_k(\boldsymbol{D}^T\otimes\boldsymbol{E})=-1+\lambda_i(\boldsymbol{D})\lambda_j(\boldsymbol{E})$$

$$k=1,2,\cdots,nm;\ i=1,2,\cdots,n;\ j=1,2,\cdots,m.$$

则是显然的.

证明完毕

特别是，在这里若设 $m=n$, $\boldsymbol{D}=\boldsymbol{A}$, $\boldsymbol{E}=\boldsymbol{A}^T$ 和 $\boldsymbol{F}=\boldsymbol{Q}$(而 $\boldsymbol{Q}=\boldsymbol{Q}^T$)，则(16)式变成

$$-\boldsymbol{X}+\boldsymbol{A}^T\boldsymbol{XA}=-\boldsymbol{Q}. \tag{17}$$

关于(17)式的解 $\boldsymbol{X}$，可以证明也有和定理 1 同样的性质.

[定理 2] 若 $\boldsymbol{A}$ 是离散系统的渐近稳定转移矩阵(即 $|\lambda_i(\boldsymbol{A})|<1$, $i=1,2,\cdots,n$)，则(17)式具有唯一解 $\boldsymbol{X}$，而且关于 $\boldsymbol{X}$ 有如下性质.

(i) $\boldsymbol{X}$ 可用

$$\boldsymbol{X}=\sum_{i=0}^{\infty}\boldsymbol{A}^{iT}\boldsymbol{Q}\boldsymbol{A}^i$$

给出，而且是有界的.

(ii)～(vi) 和定理 1(ii)～(vi)相同的性质成立.

(问题 1) 试证明定理 2.

现考察下列两种类型的线性方程式

$$\boldsymbol{XA}+\boldsymbol{A}^T\boldsymbol{X}=-\boldsymbol{Q}, \tag{18}$$

$$-\hat{\boldsymbol{X}}+\hat{\boldsymbol{A}}^T\hat{\boldsymbol{X}}\hat{\boldsymbol{A}}=-\hat{\boldsymbol{Q}}, \tag{19}$$

可见，利用下列变换

$$\left.\begin{array}{ll}\hat{\boldsymbol{A}}=(\boldsymbol{I}-\boldsymbol{A})^{-1}(\boldsymbol{I}+\boldsymbol{A}) & (\boldsymbol{A}=(\hat{\boldsymbol{A}}+\boldsymbol{I})^{-1}(\hat{\boldsymbol{A}}-\boldsymbol{I}))\\ \hat{\boldsymbol{Q}}=\boldsymbol{Q} & (\boldsymbol{Q}=\hat{\boldsymbol{Q}})\\ \hat{\boldsymbol{X}}=\dfrac{1}{2}(\boldsymbol{I}-\boldsymbol{A}^T)\boldsymbol{X}(\boldsymbol{I}-\boldsymbol{A}) & (\boldsymbol{X}=2(\hat{\boldsymbol{A}}^T+\boldsymbol{I})^{-1}\hat{\boldsymbol{X}}(\hat{\boldsymbol{A}}+\boldsymbol{I})^{-1})\end{array}\right\} \tag{20}$$

(其中假定$(\boldsymbol{I}-\boldsymbol{A})$和$(\hat{\boldsymbol{A}}+\boldsymbol{I})$均为正则的)

必定可以将其中的一种类型变换成另外一种. 而且也容易证明,对于(20)式关系中的 $\boldsymbol{A}$, $\boldsymbol{X}$, $\hat{\boldsymbol{A}}$, $\hat{\boldsymbol{X}}$,下列等值关系成立.

a) $\mathrm{Re}\{\lambda_i(\boldsymbol{A})\}<0 \quad i=1, 2, \cdots, n$

$\Leftrightarrow|\lambda_i(\hat{\boldsymbol{A}})|<1 \quad i=1, 2, \cdots, n,$

b) $\boldsymbol{X}$ 是(18)式的解 $\Leftrightarrow \hat{\boldsymbol{X}}$ 是(19)式的解.

这个结果表明,上述两种类型的线性方程式,可以利用(20)式变换将其中任何一种类型的求解问题归结到另外一种类型.

其 2 黎卡提型非线性方程式

2.1 微分方程式

现在考虑对于给定矩阵 $\boldsymbol{A}\in R^{n\times n}$, $\boldsymbol{B}\in R^{n\times m}$, $\boldsymbol{Q}=\boldsymbol{Q}^T\in R^{n\times n}$, $\boldsymbol{P}_0=\boldsymbol{P}_0^T\in R^{n\times n}$ 的非线性微分方程式[1)]

$$-\dot{\boldsymbol{P}}=\boldsymbol{PA}+\boldsymbol{A}^T\boldsymbol{P}-\boldsymbol{PBB}^T\boldsymbol{P}+\boldsymbol{Q},\ \boldsymbol{P}(0)=\boldsymbol{P}_0. \tag{21}$$

而且考察该微分方程式解的存在性. 因为该方程式是非线性的,所以满足给定起始条件解的存在不自明,而且既使存在,它也不是对于任意 $t\in R$ 都存在(例如,微分方程式 $\dot{p}=p^2+1$, $p(0)=0$ 的解 $p(t)=\tan t$ 在 $t=\frac{1}{2}\pi$ 处不存在,变成 $+\infty$). 因此,先假定解存在,研究一下它具有什么性质. 以下我们将微分方程式(21)的解表示成 $\boldsymbol{P}(t; 0, \boldsymbol{P}_0)$,或简单地写成 $\boldsymbol{P}(t)$.

首先,假定 $\boldsymbol{Q}$, $\boldsymbol{P}_0$ 为对称的,由此得下列引理.

[引理 5] 若微分方程式(21)的解 $\boldsymbol{P}(t; 0, \boldsymbol{P}_0)$ 在 $t_0\leqslant t\leqslant 0$ 存在,则具有下列性质.

(i) $\boldsymbol{P}(t; 0, \boldsymbol{P}_0)=\boldsymbol{P}(t; 0, \boldsymbol{P}_0)^T \quad t_0\leqslant t\leqslant 0$

(ii) $\boldsymbol{P}_0\geqslant\boldsymbol{0}(\boldsymbol{P}_0>\boldsymbol{0})$ 且 $\boldsymbol{Q}\geqslant\boldsymbol{0}\Rightarrow\boldsymbol{P}(t; 0, \boldsymbol{P}_0)\geqslant\boldsymbol{0}(>\boldsymbol{0}) \quad t_0\leqslant t\leqslant 0$

(iii) $\boldsymbol{P}_0\leqslant\boldsymbol{0}(\boldsymbol{P}_0<\boldsymbol{0})$ 且 $\boldsymbol{Q}\leqslant\boldsymbol{0}\Rightarrow\boldsymbol{P}(t; 0, \boldsymbol{P}_0)\leqslant\boldsymbol{0}(<\boldsymbol{0}) \quad t_0\leqslant t\leqslant 0$

(iv) $\boldsymbol{Q}\geqslant\boldsymbol{0}(\leqslant\boldsymbol{0})\Rightarrow\boldsymbol{P}(t_1; 0,\boldsymbol{0})\geqslant\boldsymbol{P}(t_2; 0,\boldsymbol{0}) \quad (\boldsymbol{P}(t_1; 0,\boldsymbol{0})\leqslant\boldsymbol{P}(t_2; 0, \boldsymbol{0}) \quad t_0\leqslant t_1\leqslant t_2\leqslant 0$

(v) 最优调节器问题

$$\dot{\boldsymbol{x}}=\boldsymbol{Ax}+\boldsymbol{Bu} \tag{22a}$$

$$J(t, \boldsymbol{u}, \boldsymbol{P}_0)=\int_t^0(\boldsymbol{x}^T\boldsymbol{Qx}+\boldsymbol{u}^T\boldsymbol{u})dt+\boldsymbol{x}(0)^T\boldsymbol{P}_0\boldsymbol{x}(0) \quad t_0\leqslant t\leqslant 0 \tag{22b}$$

的最小性能指标可以用下式给出

$$\min_u J(t, \boldsymbol{u}, \boldsymbol{P}_0)=\boldsymbol{x}(t)^T\boldsymbol{P}(t; 0, \boldsymbol{P}_0)\boldsymbol{x}(t) \tag{23}$$

(vi) $\boldsymbol{P}_0^1\geqslant\boldsymbol{P}_0^2\Rightarrow\boldsymbol{P}(t; 0, \boldsymbol{P}_0^1)\geqslant\boldsymbol{P}(t; 0, \boldsymbol{P}_0^2) \quad t_0\leqslant t\leqslant 0$

(vii) 设 $\boldsymbol{Q}=\boldsymbol{Q}_i(i=1, 2)$ 时的解为 $\boldsymbol{P}_i(t; 0, \boldsymbol{P}_0)$,则

$$\boldsymbol{Q}_1\geqslant\boldsymbol{Q}_2\Rightarrow\boldsymbol{P}_1(t; 0, \boldsymbol{P}_0)\geqslant\boldsymbol{P}_2(t; 0, \boldsymbol{P}_0) \quad t_0\leqslant t\leqslant 0$$

(viii) 设 $\boldsymbol{K}\in R^{m\times n}$ 为任意常数矩阵, $\tilde{\boldsymbol{P}}(t; 0, \boldsymbol{P}_0)$ 为线性微分方程式

$$-\dot{\tilde{\boldsymbol{P}}}=\tilde{\boldsymbol{P}}(\boldsymbol{A}+\boldsymbol{BK})+(\boldsymbol{A}+\boldsymbol{BK})^T\tilde{\boldsymbol{P}}+\boldsymbol{Q}+\boldsymbol{K}^T\boldsymbol{K},\ \tilde{\boldsymbol{P}}(0)=\boldsymbol{P}_0 \tag{24}$$

的解,则

---

1) 也可以考虑更一般的 $-\dot{\boldsymbol{P}}=\boldsymbol{PD}(t)+\boldsymbol{E}(t)\boldsymbol{P}+\boldsymbol{F}(t)-\boldsymbol{PG}(t)\boldsymbol{P}$, 但此处只限于考虑这种应用上更为重要的类型. 这种方程式也可以看成是纯量微分方程式 $\dot{p}=ap+bp^2+q$ 的推广,故称之为黎卡提型(Riccati type).

$$\boldsymbol{P}(t;0,\ \boldsymbol{P}_0)\leqslant\widetilde{\boldsymbol{P}}(t;0,\ \boldsymbol{P}_0)\quad t_0\leqslant t\leqslant 0 \tag{25}$$

成立.

（证明）

(i)～(iv) 因(21)式变化一下可表示成下列形式

$$-\dot{\boldsymbol{P}}=\boldsymbol{P}\left(\boldsymbol{A}-\frac{1}{2}\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}\right)+\left(\boldsymbol{A}-\frac{1}{2}\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}\right)^T\boldsymbol{P}+\boldsymbol{Q},$$

$$\boldsymbol{P}(0)=\boldsymbol{P}_0,$$

由引理1,若解 $\boldsymbol{P}(t)$ 存在,则可用下式给出

$$\boldsymbol{P}(t;0,\ \boldsymbol{P}_0)=\boldsymbol{\Phi}(0,\ t)^T\boldsymbol{P}_0\boldsymbol{\Phi}(0,\ t)+\int_t^0\boldsymbol{\Phi}(\tau,\ t)^T\boldsymbol{Q}\boldsymbol{\Phi}(\tau,\ t)d\tau.$$

式中 $\boldsymbol{\Phi}(0,\ t)$ 是用

$$\frac{\partial}{\partial t}\boldsymbol{\Phi}(0,\ t)=-\boldsymbol{\Phi}(0,\ t)\left(\boldsymbol{A}-\frac{1}{2}\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}(t)\right),$$

$$\boldsymbol{\Phi}(0,\ 0)=\boldsymbol{I}_n$$

定义的正则矩阵. 利用这个关系很容易得到性质(i)～(iv).

(v) 由(21), (22b)式得

$$\begin{aligned}J(t,\ \boldsymbol{u},\ \boldsymbol{P}_0)&=\int_t^0\left(\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x}+\boldsymbol{u}^T\boldsymbol{u}+\frac{d}{dt}\boldsymbol{x}^T\boldsymbol{P}\boldsymbol{x}\right)dt\\&\quad+\boldsymbol{x}(0)^T[\boldsymbol{P}_0-\boldsymbol{P}(0)]\boldsymbol{x}(0)+\boldsymbol{x}(t)^T\boldsymbol{P}(t)\boldsymbol{x}(t)\\&=\int_t^0(\boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x}+\boldsymbol{u}^T\boldsymbol{u}+\dot{\boldsymbol{x}}^T\boldsymbol{P}\boldsymbol{x}+\boldsymbol{x}^T\dot{\boldsymbol{P}}\boldsymbol{x}+\boldsymbol{x}^T\boldsymbol{P}\dot{\boldsymbol{x}})dt+\boldsymbol{x}(t)^T\boldsymbol{P}(t)\boldsymbol{x}(t),\end{aligned}$$

将(21), (22a)式代入上式,整理后变成

$$J(t,\ \boldsymbol{u},\ \boldsymbol{P}_0)=\int_t^0\|\boldsymbol{u}+\boldsymbol{B}^T\boldsymbol{P}\boldsymbol{x}\|^2dt+\boldsymbol{x}(t)^T\boldsymbol{P}(t)\boldsymbol{x}(t)\geqslant\boldsymbol{x}(t)^T\boldsymbol{P}(t)\boldsymbol{x}(t),$$

于是得到(23)式.

(vi) 根据(22b)式,由(23)式显然

$$\begin{aligned}\boldsymbol{P}_0^1\geqslant\boldsymbol{P}_0^2&\Rightarrow J(t,\ \boldsymbol{u},\ \boldsymbol{P}_0^1)\geqslant J(t,\ \boldsymbol{u},\ \boldsymbol{P}_0^2)\quad \forall\boldsymbol{u}\\&\Rightarrow\min_{\boldsymbol{u}}J(t,\ \boldsymbol{u},\ \boldsymbol{P}_0^1)\geqslant\min_{\boldsymbol{u}}J(t,\ \boldsymbol{u},\ \boldsymbol{P}_0^2)\end{aligned}$$

成立.

(vii) 证法与(vi)相同.

(viii) 在(v)的最优调节器问题中,特别是对于 $\boldsymbol{u}$ 若用 $\tilde{\boldsymbol{u}}\triangleq\boldsymbol{K}\boldsymbol{x}$, 则和(v)的证法一样,可以证明

$$J(t,\ \tilde{\boldsymbol{u}},\ \boldsymbol{P}_0)=\boldsymbol{x}(t)^T\widetilde{\boldsymbol{P}}(t;0,\ \boldsymbol{P}_0)\boldsymbol{x}(t)\quad\forall\boldsymbol{x}(t)\in R^{n\times n}.$$

由此及(23)式显然.

证明完毕

以下为了方便起见,在微分方程式(21)中特别限定 $\boldsymbol{Q}=\boldsymbol{C}^T\boldsymbol{C}$ 或 $\boldsymbol{Q}=-\boldsymbol{C}^T\boldsymbol{C}$. 以上面结果为基础,我们来考察下列微分方程式

$$-\dot{\boldsymbol{P}}=\boldsymbol{P}\boldsymbol{A}+\boldsymbol{A}^T\boldsymbol{P}-\boldsymbol{P}\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}+\boldsymbol{C}^T\boldsymbol{C},\ \boldsymbol{P}(0)=\boldsymbol{P}_0 \tag{26}$$

$$-\dot{\boldsymbol{P}}=\boldsymbol{P}\boldsymbol{A}+\boldsymbol{A}^T\boldsymbol{P}-\boldsymbol{P}\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}-\boldsymbol{C}^T\boldsymbol{C},\ \boldsymbol{P}(0)=\boldsymbol{P}_0 \tag{27}$$

解的存在性.

首先，有下列定理.

[定理 3] 若$(\boldsymbol{A}, \boldsymbol{B})$可稳定，则对于任意$\boldsymbol{P}_0\geqslant\boldsymbol{0}(\boldsymbol{P}_0>\boldsymbol{0})$，微分方程式(26)的唯一解$\boldsymbol{P}(t;0, \boldsymbol{P}_0)$在$-\infty<t\leqslant 0$存在，对于$t\leqslant 0$一致有界，而且是实对称准正定(正定)的.

(证明) 因微分方程式(26)满足局部利普希茨条件，若$|t|$足够小，则对于任意$\boldsymbol{P}_0$唯一解$\boldsymbol{P}(t;0, \boldsymbol{P}_0)$存在. 其次，对于任意$t_0<0$，若$\boldsymbol{P}(t;0, \boldsymbol{P}_0)$在区间$t_0\leqslant t\leqslant 0$存在，则表明其对于$t_0\leqslant t\leqslant 0$一致有界. 于是可以说，$\boldsymbol{P}(t;0, \boldsymbol{P}_0)$在$-\infty<t\leqslant 0$存在，而且对于$t\leqslant 0$一致有界.

首先，根据引理5(ii)，显然$\boldsymbol{P}(t;0, \boldsymbol{P}_0)$一致下有界. 而且，在引理5(viii)中，对于$\boldsymbol{K}$若选择成使$\boldsymbol{A}+\boldsymbol{BK}$渐近稳定(根据$(\boldsymbol{A}, \boldsymbol{B})$可稳定的假定，这是可能的)，由系1，系2，则微分方程式(24)的实对称解$\widetilde{\boldsymbol{P}}(t;0, \boldsymbol{P}_0)$在$-\infty<t\leqslant 0$存在，而且对于$t\leqslant 0$一致有界，由此及(25)式显然$\boldsymbol{P}(t;0, \boldsymbol{P}_0)$一致上有界. 因此，$\boldsymbol{P}(t;0, \boldsymbol{P}_0)$在$-\infty<t\leqslant 0$存在，对于$t\leqslant 0$一致有界，而且是实矩阵.

最后，由引理5(i)，(ii)，显然$\boldsymbol{P}(t;0, \boldsymbol{P}_0)$是对称准正定或正定的.

证明完毕

下面，在讨论微分方程式(27)解的存在性之前，先来讨论一下将它稍微变化一下的下列微分方程式

$$-\dot{\boldsymbol{P}}=\boldsymbol{PA}+\boldsymbol{A}^T\boldsymbol{P}-\boldsymbol{PBB}^T\boldsymbol{P}+\alpha\boldsymbol{C}^T\boldsymbol{C},\ \boldsymbol{P}(0)=\boldsymbol{0}\quad(-1<\alpha<0) \tag{28}$$

解的性质.

[引理 6] 假定$\boldsymbol{A}$渐近稳定，而且

$$\hat{\boldsymbol{H}}(-i\omega, i\omega)\triangleq\boldsymbol{I}-\boldsymbol{B}^T(-i\omega\boldsymbol{I}-\boldsymbol{A}^T)^{-1}\boldsymbol{C}^T\boldsymbol{C}(i\omega\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}\geqslant\boldsymbol{0} \tag{29}$$

几乎对于所有的实数ω都成立. 这时，若微分方程式(28)的解$\boldsymbol{P}_\alpha(t)$在$t_0\leqslant t\leqslant 0$存在，则

$$\begin{aligned}-\infty<\frac{\alpha}{1+\alpha}\int_0^\infty e^{A^T\tau}\boldsymbol{C}^T\boldsymbol{C}e^{A\tau}d\tau\leqslant\boldsymbol{P}_\alpha(t)\\ \leqslant\alpha\int_0^{-t}e^{A^T\tau}\boldsymbol{C}^T\boldsymbol{C}e^{A\tau}d\tau\leqslant\boldsymbol{0}\quad t_0\leqslant t\leqslant 0\quad -1<\alpha<0.\end{aligned} \tag{30}$$

(证明) 首先考虑下列最优调节器问题，

$$\dot{\boldsymbol{x}}=\boldsymbol{Ax}+\boldsymbol{Bu},\ \boldsymbol{y}=\boldsymbol{Cx},\ \boldsymbol{x}(t)=\boldsymbol{x}_0 \tag{31a}$$

$$J_\alpha(t, \boldsymbol{u})=\int_t^0(\alpha\boldsymbol{y}^T\boldsymbol{y}+\boldsymbol{u}^T\boldsymbol{u})d\tau\quad(t_0\leqslant t\leqslant 0). \tag{31b}$$

根据引理5(v)，则最小性能指标可用

$$\min_u J_\alpha(t, \boldsymbol{u})=\boldsymbol{x}_0^T\boldsymbol{P}_\alpha(t)\boldsymbol{x}_0\quad\forall\boldsymbol{x}_0\in R^n \tag{32}$$

给出. 而且，将系统(31a)的输出表示成

$$\boldsymbol{y}(\tau)=\boldsymbol{y}_1(\tau)+\boldsymbol{y}_2(\tau) \tag{33a}$$

$$\boldsymbol{y}_1(\tau)\triangleq\boldsymbol{C}e^{A(\tau-t)}\boldsymbol{x}_0 \tag{33b}$$

$$\boldsymbol{y}_2(\tau)\triangleq\int_t^\tau\boldsymbol{C}e^{A(\tau-\sigma)}\boldsymbol{Bu}(\sigma)d\sigma, \tag{33c}$$

代入(31b)式，得

$$J_\alpha(t, \boldsymbol{u})=J_1+J_2 \tag{34a}$$

$$J_1\triangleq\int_t^0(\|\boldsymbol{u}\|^2-\|\boldsymbol{y}_2\|^2)d\tau \tag{34b}$$

$$J_2 \triangleq \int_t^0 [(\alpha+1)\|\boldsymbol{y}\|^2 - 2\boldsymbol{y}^T\boldsymbol{y}_1 + \|\boldsymbol{y}_1\|^2] d\tau. \tag{34c}$$

在这里,对于任意的 $\boldsymbol{u}$ 考虑控制

$$\hat{\boldsymbol{u}}(\tau) \triangleq \begin{cases} \boldsymbol{u}(\tau) & t \leqslant \tau < 0 \\ \boldsymbol{0} & \tau < t,\ 0 \leqslant \tau. \end{cases}$$

将 $\tilde{\boldsymbol{u}}$ 和 $\boldsymbol{y}_2$ 的傅里叶变换分别表示成 $\hat{\boldsymbol{u}}(i\omega)$, $\hat{\boldsymbol{y}}_2(i\omega)$, 对(34b)式应用帕塞瓦尔(Parseval)等式,得

$$\begin{aligned} J_1 &= \int_{-\infty}^{\infty} (\|\tilde{\boldsymbol{u}}\|^2 - \|\boldsymbol{y}_2\|^2) d\tau \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} [\hat{\boldsymbol{u}}^T(-i\omega)\hat{\boldsymbol{u}}(i\omega) - \hat{\boldsymbol{y}}_2^T(-i\omega)\hat{\boldsymbol{y}}_2(i\omega)] d\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{\boldsymbol{u}}^T(-i\omega)\hat{\boldsymbol{H}}(-i\omega,\ i\omega)\hat{\boldsymbol{u}}(i\omega) d\omega \geqslant 0 \end{aligned}$$

而且,根据许瓦尔兹不等式(参照 B-I-9 中(14)式),由(34c)得

$$\begin{aligned} J_2 &\geqslant \int_t^0 [(\alpha+1)\|\boldsymbol{y}\|^2 - 2\|\boldsymbol{y}\|\|\boldsymbol{y}_1\| + \|\boldsymbol{y}_1\|^2] d\tau \\ &= \int_t^0 \left[(\alpha+1)\left(\|\boldsymbol{y}\| - \frac{\|\boldsymbol{y}_1\|^2}{1+\alpha}\right)^2 + \frac{\alpha}{1+\alpha}\|\boldsymbol{y}_1\|^2\right] d\tau \\ &\geqslant \frac{\alpha}{1+\alpha}\int_t^0 \|\boldsymbol{y}_1\|^2 d\tau \quad (\because\ \alpha+1>0) \\ &\geqslant \left(\frac{\alpha}{1+\alpha}\right)\boldsymbol{x}_0^T\left[\int_0^{\infty} e^{A^T\tau}\boldsymbol{C}^T\boldsymbol{C}e^{A\tau} d\tau\right]\boldsymbol{x}_0 \quad \left(\because \frac{\alpha}{1+\alpha}<0\right) \\ &> -\infty \quad (\text{根据定理 1 的(i)}) \end{aligned}$$

因此,对于任意 $\boldsymbol{u}$,

$$J_\alpha(t,\ \boldsymbol{u}) = J_1 + J_2 \geqslant \frac{\alpha}{1+\alpha}\boldsymbol{x}_0^T\left[\int_0^{\infty} e^{A^T\tau}\boldsymbol{C}^T\boldsymbol{C}e^{A\tau} d\tau\right]\boldsymbol{x}_0 > -\infty \quad \forall\boldsymbol{u},\ \forall\boldsymbol{x}_0 \in \boldsymbol{R}^n \tag{35}$$

成立. 而且,在(31b),(33)式中若令 $\boldsymbol{u}=\boldsymbol{0}$, 可得

$$J_\alpha(t,\ \boldsymbol{0}) = \alpha\int_t^0 \|\boldsymbol{y}_1\|^2 dt = \alpha\boldsymbol{x}_0^T\left[\int_0^{-t} e^{A^T\tau}\boldsymbol{C}^T\boldsymbol{C}e^{A\tau} d\tau\right]\boldsymbol{x}_0 \leqslant 0 \quad \forall\boldsymbol{x}_0 \in R^n. \tag{36}$$

由(32),(35),(36)式可得(30)式.

证明完毕

由上述结果立即可以得出微分方程式(28)的解存在的条件.

[定理 4] 若 $\boldsymbol{A}$ 渐近稳定且满足(29)式,则对于满足 $-1<\alpha<0$ 的任意 α, (28)式的唯一解 $\boldsymbol{P}_\alpha(t)$ 在 $-\infty<t\leqslant 0$ 存在,对于 $t\leqslant 0$ 一致有界,而且是实对称准负定的.

(证明) 利用引理 6 的结果进行和定理 3 同样的证明即可.

证明完毕

2.2 代数方程式

本节将利用前节的结果考察与微分方程式(21)对应的下列代数方程式

$$\boldsymbol{PA} + \boldsymbol{A}^T\boldsymbol{P} - \boldsymbol{PBB}^T\boldsymbol{P} + \boldsymbol{Q} = \boldsymbol{0} \tag{37}$$

解的存在性. 式中对于矩阵 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{Q}$ 的假定和前节一样. 因该方程式是非线性的, 由二次方程式的例子也可以理解, 一般即使解存在, 也不一定存在实对称解. 在这里我们只限于考察微分方程式的平衡点, 即只考虑实对称解.

首先, 假定(37)式的实对称解存在, 来研究解的性质.

[引理 7]

a) 假定代数方程式(37)存在两个相异的实对称解 $\boldsymbol{P}_1$ 和 $\boldsymbol{P}_2$. 这时, 若$(\boldsymbol{A}, \boldsymbol{B})$可控, 则有下列性质.

i) $\boldsymbol{P}_1 > \boldsymbol{P}_2 \Rightarrow \mathrm{Re}\,\lambda(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}_1) < 0,$

$\qquad \mathrm{Re}\,\lambda(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}_2) > 0.$

b) 反之, 假定在(37)式的实对称解中有满足

$$\mathrm{Re}\,\lambda(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}) < 0 \quad (>0) \tag{38}$$

的解存在, 将其表示成 $\boldsymbol{P}_+(\boldsymbol{P}_-)$, 则有下列性质.

ii) $\boldsymbol{P}_+$ 和 $\boldsymbol{P}_-$ 唯一确定.

iii) 当$(\boldsymbol{A}, \boldsymbol{B})$可控时, 若 $\boldsymbol{P}_+$ 和 $\boldsymbol{P}_-$ 之中有任意一个存在, 则另外一个也必定存在, 而且

$$\boldsymbol{P}_+ - \boldsymbol{P}_- = \left(\int_0^\infty e^{\boldsymbol{A}_+ t}\boldsymbol{B}\boldsymbol{B}^T e^{(\boldsymbol{A}_+)^T t}\,dt\right)^{-1} \tag{39a}$$

$$= \left(\int_{-\infty}^0 e^{\boldsymbol{A}_- t}\boldsymbol{B}\boldsymbol{B}^T e^{(\boldsymbol{A}_-)^T t}\,dt\right)^{-1} \tag{39b}$$

成立. 其中设

$$\boldsymbol{A}_+ \triangleq \boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}_+,\ \boldsymbol{A}_- \triangleq \boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}_-.$$

iv) 设在(37)式中令 $\boldsymbol{Q}=\tilde{\boldsymbol{Q}}$ 时的任意实对称解为 $\tilde{\boldsymbol{P}}$, 则

$$\boldsymbol{Q} \geqslant \tilde{\boldsymbol{Q}}\,(\boldsymbol{Q} > \tilde{\boldsymbol{Q}}) \Rightarrow \boldsymbol{P}_+ \geqslant \tilde{\boldsymbol{P}} \geqslant \boldsymbol{P}_-\,(\boldsymbol{P}_+ > \tilde{\boldsymbol{P}} > \boldsymbol{P}_-)$$

成立.

(证明)

i) 令 $\Delta\boldsymbol{P} \triangleq \boldsymbol{P}_1 - \boldsymbol{P}_2$, 由(37)式容易得

$$\Delta\boldsymbol{P}(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}_1) + (\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}_1)^T\Delta\boldsymbol{P} = -\Delta\boldsymbol{P}\boldsymbol{B}\boldsymbol{B}^T\Delta\boldsymbol{P}.$$

设 $\Delta\boldsymbol{P} > \boldsymbol{0}$, 在上式分别左、右乘以$(\Delta\boldsymbol{P})^{-1}$, 则

$$(\Delta\boldsymbol{P})^{-1}(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}_1)^T + (\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}_1)(\Delta\boldsymbol{P})^{-1} = -\boldsymbol{B}\boldsymbol{B}^T.$$

根据假定, 这里因$(\boldsymbol{B}^T, (\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}_1)^T)$可观测, 由定理 1(ix), 则 $\mathrm{Re}\,\lambda(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}_1) < 0.$

设 $\Delta\boldsymbol{P} \triangleq \boldsymbol{P}_2 - \boldsymbol{P}_1$, 进行完全同样的讨论可得 $\mathrm{Re}\,\lambda(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}_2) > 0.$

ii) 设 $\boldsymbol{P}_1, \boldsymbol{P}_2$ 为(37)式的任意实对称解, 则容易得

$$(\boldsymbol{P}_1-\boldsymbol{P}_2)(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}_1) + (\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}_2)^T(\boldsymbol{P}_1-\boldsymbol{P}_2) = \boldsymbol{0},$$

在这里若假定 $\boldsymbol{P}_1, \boldsymbol{P}_2$ 都满足(38)式, 根据引理 2 由 $\boldsymbol{P}_1-\boldsymbol{P}_2=\boldsymbol{0}$ 显然.

iii) 设 $\boldsymbol{P}_+$ 存在, 因 $\boldsymbol{P}_+$ 满足(38)式, 由定理 1, 则方程式

$$\boldsymbol{X}(\boldsymbol{A}_+)^T + \boldsymbol{A}_+\boldsymbol{X} = -\boldsymbol{B}\boldsymbol{B}^T$$

具有唯一解

$$\boldsymbol{X} = \int_0^\infty e^{\boldsymbol{A}_+ t}\boldsymbol{B}\boldsymbol{B}^T e^{(\boldsymbol{A}_+)^T t}\,dt \tag{40}$$

由于假定$(\boldsymbol{A}, \boldsymbol{B})$可控, 则 $\boldsymbol{X} > 0$. 因此, 若令 $\Delta\boldsymbol{P} \triangleq \boldsymbol{X}^{-1}$, 显然 $\Delta\boldsymbol{P}$ 满足

$$(\boldsymbol{A}_+)^T\Delta\boldsymbol{P}+\Delta\boldsymbol{P}\boldsymbol{A}_+=-\Delta\boldsymbol{P}\boldsymbol{B}\boldsymbol{B}^T\Delta\boldsymbol{P}.$$

将该式和(37)式对比，则可很简单地证明 $\boldsymbol{P}\triangleq\boldsymbol{P}_+-\Delta\boldsymbol{P}$ 是(37)式的实对称解. 但因 $\boldsymbol{P}_+-\boldsymbol{P}=\Delta\boldsymbol{P}>\boldsymbol{0}$，将前面的性质 i)，ii) 合起来，则可见 $\boldsymbol{P}=\boldsymbol{P}_-$. 由此及(40)式可得(39a)式.

反之，设 $\boldsymbol{P}_-$ 存在，完全同样可以证明 $\boldsymbol{P}_+$ 存在，而且(39b)式成立.

iv) 设 $\Delta\boldsymbol{P}\triangleq\boldsymbol{P}_+-\widetilde{\boldsymbol{P}}$，$\Delta\boldsymbol{Q}=\boldsymbol{Q}-\widetilde{\boldsymbol{Q}}$，由(37)式得

$$\Delta\boldsymbol{P}\boldsymbol{A}_++(\boldsymbol{A}_+)^T\Delta\boldsymbol{P}=-\Delta\boldsymbol{P}\boldsymbol{B}\boldsymbol{B}^T\Delta\boldsymbol{P}-\Delta\boldsymbol{Q}.$$

注意，该式右边是准负定(负定)的，应用定理 1(iii)，(ii)，得 $\Delta\boldsymbol{P}\geqslant\boldsymbol{0}(\Delta\boldsymbol{P}>\boldsymbol{0})$. 对于 $\boldsymbol{P}_-$ 也是同样.

证明完毕

(注意)

i) 在性质 iv) 中特别是若考虑 $\widetilde{\boldsymbol{Q}}=\boldsymbol{Q}$ 的情况，可见 $\boldsymbol{P}_+(\boldsymbol{P}_-)$ 是(37)式实对称解中最大(最小)的解. 因此，称它为(37)式的最大解(最小解). 再由性质 iv) 可见，最大解(最小解)将随着 $\boldsymbol{Q}$ 的增加单调非减少(单调非增大).

ii) 假定除了 $\boldsymbol{P}_+$，$\boldsymbol{P}_-$ 以外还有实对称解 $\boldsymbol{P}$ 存在，则 $\boldsymbol{P}_+\geqslant\boldsymbol{P}\geqslant\boldsymbol{P}_-$ 成立. 但是当 $(\boldsymbol{A},\ \boldsymbol{B})$ 可控时，$\boldsymbol{P}_+>\boldsymbol{P}$ 或 $\boldsymbol{P}>\boldsymbol{P}_-$ 不成立. 其原因是，如假定 $\boldsymbol{P}_+>\boldsymbol{P}(\boldsymbol{P}>\boldsymbol{P}_-)$，由 i) 则 $\mathrm{Re}\,\lambda(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P})>0(<0)$，由 ii) 则必须 $\boldsymbol{P}=\boldsymbol{P}_-(\boldsymbol{P}=\boldsymbol{P}_+)$.

以下为了方便起见，在(37)式中只限定 $\boldsymbol{Q}=\boldsymbol{C}^T\boldsymbol{C}$ 或 $\boldsymbol{Q}=-\boldsymbol{C}^T\boldsymbol{C}$，先考察下列两个方程式

$$\boldsymbol{P}\boldsymbol{A}+\boldsymbol{A}^T\boldsymbol{P}-\boldsymbol{P}\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}+\boldsymbol{C}^T\boldsymbol{C}=\boldsymbol{0} \tag{41}$$

$$\boldsymbol{P}\boldsymbol{A}+\boldsymbol{A}^T\boldsymbol{P}-\boldsymbol{P}\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}-\boldsymbol{C}^T\boldsymbol{C}=\boldsymbol{0}. \tag{42}$$

(在(42)式中假定 $\boldsymbol{C}$ 是正则的)

以后将其结果推广到一般的 $\boldsymbol{Q}=\boldsymbol{Q}^T$ 的情况.

首先，有下面定理.

[定理 5]

i) 若 $(\boldsymbol{A},\ \boldsymbol{B})$ 可稳定，则代数方程式(41)的实对称解中至少有一个准正定解存在.

ii) 此外若再假定 $(\boldsymbol{C},\ \boldsymbol{A})$ 可检测(可观测)，则准正定解仅限一个，而且若设其为 $\boldsymbol{P}$，则

$$\mathrm{Re}\,\lambda(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P})<0 \tag{43}$$

(及 $\boldsymbol{P}>\boldsymbol{0}$) 成立.

(证明)

i) 根据定理 3，微分方程式(26)的解 $\boldsymbol{P}(t;\,0,\ \boldsymbol{0})$ 在 $-\infty<t\leqslant0$ 存在，对于 $t\leqslant0$ 一致有界而且是实对称准正定的. 此外，根据引理 5(iv)，随着 t 的减少 $\boldsymbol{P}(t;\,0,\ \boldsymbol{0})$ 还是单调非减少的，因此 $\lim\limits_{t\to-\infty}\boldsymbol{P}(t;\,0,\ \boldsymbol{0})\triangleq\boldsymbol{P}$ 存在(参照 B-I-11 定理 19)，它是(41)式的实对称准正定解.

ii) 首先来证明 i) 中得到的(41)式的准正定解 $\boldsymbol{P}$ 满足(43)式. (41)式变化后可以表示成

$$\boldsymbol{P}(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P})+(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P})^T\boldsymbol{P}+\boldsymbol{P}\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}+\boldsymbol{C}^T\boldsymbol{C}=\boldsymbol{0}. \tag{44}$$

现假定 $\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}$ 具有满足 $\mathrm{Re}\lambda\geqslant 0$ 的特征值 $\lambda\in C(\mathrm{Re}\,\lambda\geqslant 0)$, 设对应于 λ 的 $(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P})$ 的特征向量为 $\boldsymbol{x}\in C^n(\boldsymbol{x}\neq\boldsymbol{0})$. 则

$$(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P})\boldsymbol{x}=\lambda\boldsymbol{x}. \tag{45}$$

对(44)式左乘以 $\boldsymbol{x}^*$, 右乘以 $\boldsymbol{x}$, 将(45)式代入后, 得

$$2\,\mathrm{Re}\,\lambda\,\boldsymbol{x}^*\boldsymbol{P}\boldsymbol{x}+\boldsymbol{x}^*\boldsymbol{P}\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}\boldsymbol{x}+\boldsymbol{x}^*\boldsymbol{C}^T\boldsymbol{C}\boldsymbol{x}=0.$$

因 $\mathrm{Re}\,\lambda\geqslant 0$ 且 $\boldsymbol{P}\geqslant\boldsymbol{0}$, 为使该式成立则必须

$$\boldsymbol{B}^T\boldsymbol{P}\boldsymbol{x}=\boldsymbol{0} \tag{46a}$$

$$\boldsymbol{C}\boldsymbol{x}=\boldsymbol{0}. \tag{46b}$$

将(46a)式代入(45)式, 得

$$\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}\quad(\mathrm{Re}\,\lambda\geqslant 0). \tag{47}$$

显然, (46b), (47)式与可检测性相矛盾. 因此, (43)式成立.

特别是, 若 $(\boldsymbol{C},\ \boldsymbol{A})$ 可观测, 则 $\boldsymbol{P}>\boldsymbol{0}$ 可证明如下. 设

$$\boldsymbol{F}\triangleq\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P},\ \boldsymbol{D}\triangleq\begin{bmatrix}\boldsymbol{B}^T\boldsymbol{P}\\ \boldsymbol{C}\end{bmatrix},$$

则(44)式可以表示成

$$\boldsymbol{P}\boldsymbol{F}+\boldsymbol{F}^T\boldsymbol{P}+\boldsymbol{D}^T\boldsymbol{D}=\boldsymbol{0},$$

而且当令 $\lambda\in C,\ \boldsymbol{x}\in C^n$ 时

$$\boldsymbol{F}\boldsymbol{x}=\lambda\boldsymbol{x},\ \boldsymbol{D}\boldsymbol{x}=\boldsymbol{0}\Rightarrow\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x},\ \boldsymbol{C}\boldsymbol{x}=\boldsymbol{0}$$

成立. 考虑到 $(\boldsymbol{C},\ \boldsymbol{A})$ 可观测, 显然 $(\boldsymbol{D},\ \boldsymbol{F})$ 亦可观测(参照 A-I-16 附录中系 1). 因此, 根据定理 1(vi), $\boldsymbol{P}>\boldsymbol{0}$.

最后, 根据(43)式及引理 7(ii), 显然 $\boldsymbol{P}$ 是(41)式的唯一解.

证明完毕

其次, 我们来考察代数方程式(42). 在这之前先讲一下微分方程式(28)定常解的性质.

[引理 8] 假定(28)式中 $\boldsymbol{C}$ 是正则的[1]. 这时, 若 $\boldsymbol{A}$ 渐近稳定且满足(29)式, 则对于任意 $\alpha(-1<\alpha<0)$, 微分方程式(28)的定常解 $\lim\limits_{t\to-\infty}\boldsymbol{P}_\alpha(t)\triangleq\boldsymbol{P}_\alpha$ 存在, 而且具有下列性质.

i) $\boldsymbol{P}_\alpha=\boldsymbol{P}_\alpha^T\in R^{n\times n}$.

ii) $\boldsymbol{P}_\alpha<\boldsymbol{0}$

iii) $-1\leqslant\alpha_1\leqslant\alpha_2<0\Rightarrow\boldsymbol{P}_{\alpha1}\leqslant\boldsymbol{P}_{\alpha2}$

iv) $\boldsymbol{P}_\alpha\boldsymbol{A}+\boldsymbol{A}^T\boldsymbol{P}_\alpha-\boldsymbol{P}_\alpha\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}_\alpha+\alpha\boldsymbol{C}^T\boldsymbol{C}=\boldsymbol{0}$ (48)

v) $\mathrm{Re}\,\lambda(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}_\alpha)<0$ (49)

vi) 若 $(\boldsymbol{A},\ \boldsymbol{B})$ 可控, 则

$$\boldsymbol{P}_\alpha\geqslant-\left[\int_0^\infty e^{At}\boldsymbol{B}\boldsymbol{B}^T e^{A^Tt}\,dt\right]^{-1} \tag{50}$$

成立.

(证明) 根据定理 4 和引理 5(iv), 因 $\boldsymbol{P}_\alpha(t)$ 在 $-\infty<t\leqslant 0$ 存在, 对于 $t\leqslant 0$ 一致有界且随着 t 的减少单调非增大, 则极限值 $\boldsymbol{P}_\alpha$ 存在.

1) 这里为了证明方便起见假定 $\boldsymbol{C}$ 是正则的, 但是在 $\boldsymbol{C}$ 不是正则矩阵的情况下也可以证明该引理成立, 只是证明稍复杂一些.

(这时性质 ii) 变为 $\boldsymbol{P}_\alpha\leqslant\boldsymbol{0}$).

i）根据定理 4, $\boldsymbol{P}_\alpha(t)$是实对称的，由此及 $\boldsymbol{P}_\alpha$ 的定义显然.

ii）在引理 6 的(30)式中设 $t\to-\infty$, 因

$$\boldsymbol{P}_\alpha\leqslant\alpha\int_0^\infty e^{A^Tt}\boldsymbol{C}^T\boldsymbol{C}e^{At}\,dt<\boldsymbol{0}$$

成立，则显然.

iii）由引理 5(vii) 和 $\boldsymbol{P}_\alpha$ 的定义显然.

iv）因 $\boldsymbol{P}_\alpha$ 是满足(28)式的常数矩阵，则显然.

v）对于任意 $\alpha(-1<\alpha<0)$，选择满足 $-1<\alpha_1<\alpha$ 的 α_1，若定义 $\Delta\boldsymbol{P}\triangleq\boldsymbol{P}_\alpha-\boldsymbol{P}_{\alpha_1}$, $\Delta\boldsymbol{Q}=(\alpha-\alpha_1)\boldsymbol{C}^T\boldsymbol{C}$, 根据 iv)，则

$$\Delta\boldsymbol{P}(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}_\alpha)+(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}_\alpha)^T\Delta\boldsymbol{P}+\Delta\boldsymbol{P}\boldsymbol{B}\boldsymbol{B}^T\Delta\boldsymbol{P}+\Delta\boldsymbol{Q}=\boldsymbol{0} \tag{51}$$

成立. 现假定 $\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}_\alpha$ 具有满足 $\mathrm{Re}\,\lambda\geqslant0$ 的特征值 λ, 设与 λ 对应的特征向量为 $\boldsymbol{x}(\boldsymbol{x}\neq\boldsymbol{0})$，则

$$(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}_\alpha)\boldsymbol{x}=\lambda\boldsymbol{x}. \tag{52}$$

对(51)式左乘以 $\boldsymbol{x}^*$, 右乘以 $\boldsymbol{x}$, 将(52)式代入，则

$$2\,\mathrm{Re}\,\lambda\boldsymbol{x}^*\Delta\boldsymbol{P}\boldsymbol{x}+\boldsymbol{x}^*\Delta\boldsymbol{P}\boldsymbol{B}\boldsymbol{B}^T\Delta\boldsymbol{P}\boldsymbol{x}+\boldsymbol{x}^*\Delta\boldsymbol{Q}\boldsymbol{x}=0$$

成立. 因 $\mathrm{Re}\lambda\geqslant0$, $\Delta\boldsymbol{Q}>\boldsymbol{0}$, 而且根据 iii), $\Delta\boldsymbol{P}\geqslant\boldsymbol{0}$, 则为使上式成立必须 $\Delta\boldsymbol{Q}\boldsymbol{x}=\boldsymbol{0}$, 即必须 $\boldsymbol{x}=\boldsymbol{0}$, 这是矛盾的.

vi）设

$$\tilde{\boldsymbol{P}}\triangleq\int_0^\infty e^{At}\boldsymbol{B}\boldsymbol{B}^Te^{A^Tt}dt,$$

首先，由$(\boldsymbol{A}, \boldsymbol{B})$的可控性得，$\tilde{\boldsymbol{P}}>\boldsymbol{0}$, 而且由定理 1 可见

$$\tilde{\boldsymbol{P}}\boldsymbol{A}^T+\boldsymbol{A}\tilde{\boldsymbol{P}}+\boldsymbol{B}\boldsymbol{B}^T=\boldsymbol{0} \tag{53}$$

成立.

对上式左、右乘以 $\tilde{\boldsymbol{P}}^{-1}$，显然 $\hat{\boldsymbol{P}}\triangleq-\tilde{\boldsymbol{P}}^{-1}$ 是方程式

$$\boldsymbol{P}\boldsymbol{A}+\boldsymbol{A}^T\boldsymbol{P}-\boldsymbol{P}\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}=\boldsymbol{0} \tag{54}$$

的一个实对称解. 因为(54)式还具有除了 $\hat{\boldsymbol{P}}$ 以外 $\boldsymbol{0}$ 的自明解，而且 $\hat{\boldsymbol{P}}<\boldsymbol{0}$, 根据引理 7 i)，则

$$\mathrm{Re}\,\lambda(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\hat{\boldsymbol{P}})>0 \tag{55}$$

成立，即 $\hat{\boldsymbol{P}}$ 是(54)式的最小解（参看引理 7 的注意 i)）. 因此，将(54)和(48)式对照，根据引理 7 iv) 可以看到，对于(54)式的最小解 $\hat{\boldsymbol{P}}$ 和(48)式的解 $\boldsymbol{P}_\alpha$, $\boldsymbol{P}_\alpha\geqslant\hat{\boldsymbol{P}}$ 成立（在引理 7 iv) 中令 $\boldsymbol{Q}=\boldsymbol{0}$, $\tilde{\boldsymbol{Q}}=\alpha\boldsymbol{C}^T\boldsymbol{C}$）.

证明完毕

由上面结果立即可以得出(42)式实对称解存在的条件及其性质.

[定理 6] 设 $\boldsymbol{A}$ 渐近稳定且满足(29)式. 这时

i）若$(\boldsymbol{A}, \boldsymbol{B})$可控，则代数方程式(42)的实对称解中有满足下列条件的负定解 $\boldsymbol{P}_+$ 存在.

$$\mathrm{Re}\,\lambda(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}_+)\leqslant0 \tag{56}$$

ii）$\boldsymbol{P}_+$ 是(42)式的最大解.

（证明）

i）根据引理 8 iii）和 vi），则 $\lim\limits_{\alpha\to-1}\boldsymbol{P}_{\alpha}\triangleq\boldsymbol{P}_{+}$ 存在，而且根据 i），ii），iii），iv），它是(42)式的实对称负定解．在(49)式中令 $\alpha\to-1$ 可以得到(56)式．

ii）首先，根据引理 8 iv），v）及引理 7 iv），(42)式的任意实对称解 $\boldsymbol{P}$ 满足 $\boldsymbol{P}\leqslant\boldsymbol{P}_{\alpha}$．在这里若令 $\alpha\to-1$，则可以得 $\boldsymbol{P}\leqslant\lim\limits_{\alpha\to-1}\boldsymbol{P}_{\alpha}=\boldsymbol{P}_{+}$．

以以上结果为基础，下面我们一般地考察 $\boldsymbol{Q}$ 已证明过的情况．为此，如下先将 $\boldsymbol{Q}$ 表示成某两个对称正定矩阵之差

$$\boldsymbol{Q}=r\boldsymbol{I}_n-(r\boldsymbol{I}_n-\boldsymbol{Q})=\boldsymbol{C}_1^T\boldsymbol{C}_1-\boldsymbol{C}_2^T\boldsymbol{C}_2, \tag{57}$$

（设 $r>1$，$\boldsymbol{C}_1$，$\boldsymbol{C}_2$ 均为正则的）

那么，根据定理 5，若$(\boldsymbol{A},\ \boldsymbol{B})$可控，则在(41)式中令 $\boldsymbol{C}=\boldsymbol{C}_1$ 所得到的代数方程式

$$\boldsymbol{PA}+\boldsymbol{A}^T\boldsymbol{P}-\boldsymbol{PBB}^T\boldsymbol{P}+\boldsymbol{C}_1^T\boldsymbol{C}_1=\boldsymbol{0} \tag{58}$$

的实对称解中，有满足

$$\mathrm{Re}\,\lambda(\boldsymbol{A}-\boldsymbol{BB}^T\boldsymbol{P})<0 \tag{59}$$

的解存在，而且唯一确定（用 $\boldsymbol{P}_0$ 表示它）．

其次，设 $\boldsymbol{A}_1\triangleq\boldsymbol{A}-\boldsymbol{BB}^T\boldsymbol{P}_0$，考虑代数方程式

$$\boldsymbol{XA}_1+\boldsymbol{A}_1^T\boldsymbol{X}-\boldsymbol{XBB}^T\boldsymbol{X}-\boldsymbol{C}_2^T\boldsymbol{C}_2=\boldsymbol{0}. \tag{60}$$

这样容易证明，在(37)和(60)式两个代数方程式之间下列关系成立．

［引理 9］ 设$(\boldsymbol{A},\ \boldsymbol{B})$可控，则在下列意义上(37)和(60)式是等价的．

i）若(37)式具有实对称解 $\boldsymbol{P}$，则与其相对应（唯一确定）(60)式也有某个实对称解 $\boldsymbol{X}$ 存在，而且两者之间有下列关系（其逆亦成立）

$$\boldsymbol{P}-\boldsymbol{X}=\boldsymbol{P}_0$$

ii）若(37)式具有实对称解 $\boldsymbol{P}_1$，$\boldsymbol{P}_2$，且 $\boldsymbol{P}_1\geqslant\boldsymbol{P}_2$，则对于与此相应的(60)式的实对称解 $\boldsymbol{X}_1$，$\boldsymbol{X}_2$，$\boldsymbol{X}_1\geqslant\boldsymbol{X}_2$ 成立（其逆亦成立）．（因此，若(60)式的最大解 $\boldsymbol{X}^+$ 存在，则矩阵 $\boldsymbol{P}^+\triangleq\boldsymbol{P}_0+\boldsymbol{X}^+$ 是(37)式的最大解）．

iii）设与(37)式的实对称解 $\boldsymbol{P}$ 对应的(60)式的实对称解为 $\boldsymbol{X}$，则

$$\boldsymbol{A}-\boldsymbol{BB}^T\boldsymbol{P}=\boldsymbol{A}_1-\boldsymbol{BB}^T\boldsymbol{X}$$

成立．

利用引理 9 可以得到代数方程(37)式实对称解的存在条件和实对称解的性质．

［定理 7］ 设$(\boldsymbol{A},\ \boldsymbol{B})$可控，这时代数方程式(37)的实对称解存在的充分必要条件

$$\boldsymbol{H}(-i\omega,\ i\omega)\triangleq\boldsymbol{I}+\boldsymbol{B}^T(-i\omega\boldsymbol{I}-\boldsymbol{A}^T)^{-1}\boldsymbol{Q}(i\omega\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}\geqslant\boldsymbol{0} \tag{61}$$

几乎对所有实数 ω 都成立．而且，在满足该条件的情况下，

i）满足下列条件的实对称解 $\boldsymbol{P}_+$ 和 $\boldsymbol{P}_-$ 存在，

$$\mathrm{Re}\,\lambda(\boldsymbol{A}-\boldsymbol{BB}^T\boldsymbol{P}_+)\leqslant0$$
$$\mathrm{Re}\,\lambda(\boldsymbol{A}-\boldsymbol{BB}^T\boldsymbol{P}_-)\geqslant0$$

ii）$\boldsymbol{P}_+$ 是该方程式的最大解，$\boldsymbol{P}_-$ 是最小解（因此 $\boldsymbol{P}_+$ 和 $\boldsymbol{P}_-$ 唯一确定）．

（证明）（必要性）若(37)式的实对称解 $\boldsymbol{P}$ 存在，显然

$$\boldsymbol{Q}=\boldsymbol{P}(i\omega\boldsymbol{I}-\boldsymbol{A})+(-i\omega\boldsymbol{I}-\boldsymbol{A}^T)\boldsymbol{P}+\boldsymbol{PBB}^T\boldsymbol{P}\quad\forall\omega\in R$$

成立．在该式两边左乘以 $\boldsymbol{B}^T(-i\omega\boldsymbol{I}-\boldsymbol{A}^T)^{-1}$，右乘以 $(i\omega\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}$，代入(61)式右边，则

$$\boldsymbol{H}(-i\omega,\ i\omega)=[\boldsymbol{I}+\boldsymbol{B}^T\boldsymbol{P}(-i\omega\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}]^T[\boldsymbol{I}+\boldsymbol{B}^T\boldsymbol{P}(i\omega\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}]\quad \text{a. e.}^{*}\ \omega\in R \tag{62}$$

因上式右边是准正定埃尔米特矩阵(参看 B-I-11 中埃尔米特矩阵一节),则得到(61)式.

(充分性) 首先,对于 $\boldsymbol{P}_+$ 若能证明(60)式的实对称解中有满足 $\mathrm{Re}\,\lambda(\boldsymbol{A}_1-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{X}_+)\leqslant 0$ 的解 $\boldsymbol{X}_+$ 存在,而且是(60)式的最大解,根据引理9,充分性显然(显然 $\boldsymbol{P}_+=\boldsymbol{P}_0+\boldsymbol{X}_+$).而且,由于 $\boldsymbol{A}_1$ 渐近稳定,$(\boldsymbol{A}_1,\ \boldsymbol{B})$ 可控,且

$$\begin{aligned}&\boldsymbol{I}-\boldsymbol{B}^T(-i\omega\boldsymbol{I}-\boldsymbol{A}_1^T)^{-1}\boldsymbol{C}_2^T\boldsymbol{C}_2(i\omega\boldsymbol{I}-\boldsymbol{A}_1)^{-1}\boldsymbol{B}\\&\quad=[\boldsymbol{I}+\boldsymbol{B}^T\boldsymbol{P}_0(-i\omega\boldsymbol{I}-\boldsymbol{A}^T)^{-1}\boldsymbol{B}]^{-1T}\boldsymbol{H}(-i\omega,\ i\omega)\\&\qquad\times[\boldsymbol{I}+\boldsymbol{B}^T\boldsymbol{P}_0(i\omega\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}]^{-1}\geqslant\boldsymbol{0}\quad\text{a. e.}\ \omega\in R\end{aligned} \tag{63}$$

(其证明参看本章附录 A),根据定理6,这是显然的.

对于 $\boldsymbol{P}_-$,设 $\widetilde{\boldsymbol{A}}\triangleq-\boldsymbol{A}$,考虑下列代数方程式

$$\widetilde{\boldsymbol{P}}\widetilde{\boldsymbol{A}}+\widetilde{\boldsymbol{A}}^T\widetilde{\boldsymbol{P}}-\widetilde{\boldsymbol{P}}\boldsymbol{B}\boldsymbol{B}^T\widetilde{\boldsymbol{P}}+\boldsymbol{Q}=\boldsymbol{0} \tag{64}$$

那么对于该方程式,由于满足

$$\begin{aligned}\widetilde{\boldsymbol{H}}(-i\omega,\ i\omega)&\triangleq\boldsymbol{I}+\boldsymbol{B}^T(-i\omega\boldsymbol{I}-\widetilde{\boldsymbol{A}}^T)^{-1}\boldsymbol{Q}(i\omega\boldsymbol{I}-\widetilde{\boldsymbol{A}})^{-1}\boldsymbol{B}\\&=\boldsymbol{H}(i\omega,\ -i\omega)\geqslant\boldsymbol{0}\quad\text{a. e.}\ \omega\in R\end{aligned}$$

根据上述,有满足

$$\mathrm{Re}\,\lambda(\widetilde{\boldsymbol{A}}-\boldsymbol{B}\boldsymbol{B}^T\widetilde{\boldsymbol{P}}_+)\leqslant 0 \tag{65}$$

的实对称解 $\widetilde{\boldsymbol{P}}_+$ 存在,而且这是最大解.因此,若令 $\boldsymbol{P}_-\triangleq-\widetilde{\boldsymbol{P}}_+$,首先由(64)式,$\boldsymbol{P}_-$ 是(37)式的实对称解,且由(65)式,

$$\mathrm{Re}\,\lambda(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}_-)=-\mathrm{Re}\,\lambda(\widetilde{\boldsymbol{A}}-\boldsymbol{B}\boldsymbol{B}^T\widetilde{\boldsymbol{P}}_+)\geqslant 0$$

成立.最后,因 $-\boldsymbol{P}_-(=\widetilde{\boldsymbol{P}}_+)$ 是(64)式的最大解,而且对于(37)式的任意实对称解 $\boldsymbol{P}$,$-\boldsymbol{P}$ 是(60)式的实对称解,显然 $\boldsymbol{P}_-$ 是(37)式的最小解(只要最大解 $\boldsymbol{P}_+$ 及最小解 $\boldsymbol{P}_-$ 存在,唯一性显然).

证明完毕

(注意)

由该结果可见,在 $(\boldsymbol{A},\ \boldsymbol{B})$ 可控的情况下,若代数方程式(37)的实对称解存在,则其中必然存在着最大解($\boldsymbol{P}_+$)和最小解($\boldsymbol{P}_-$)(若实对称解存在,则(61)式成立,因而 $\boldsymbol{P}_+$ 和 $\boldsymbol{P}_-$ 存在).而且,$\boldsymbol{P}_+$ 和 $\boldsymbol{P}_-$ 以外的实对称解都可以表示成该二解的某种线性组合,对此可参考文献[138].

2.3 平衡点和稳定性

对于在其1中考察的线性微分方程式(2),若 $\boldsymbol{A}$ 渐近稳定,则平衡点存在,它是唯一的且大范围渐近稳定.因这里考虑的微分方程式(21)是非线性的,所以平衡点不一定存在,而且既使存在,一般也不是唯一的.如前节所证明,若存在,则必然有 $\boldsymbol{P}_+$,$\boldsymbol{P}_-$ 的某些特征存在.因此,对该特征的两个平衡点的稳定性再稍微讲一下.

[定理8] 设 $(\boldsymbol{A},\ \boldsymbol{B})$ 可控且代数方程式(37)具有实对称解(因而最大解 $\boldsymbol{P}_+$ 和最小解 $\boldsymbol{P}_-$ 存在).这时,若 $\boldsymbol{P}_+>\boldsymbol{P}_-$,则微分方程式(21)的解 $\boldsymbol{P}(t;0,\ \boldsymbol{P}_0)$ 有下列性质

* a. e. 为 almost everywhere 之缩写,意思是几乎处处,即除了测度为0的点集…….a. e. $\omega\in R$ 表示上式对于 R 中的 ω 几乎处处都成立.——译者注

i) $\lim\limits_{t\to-\infty} \boldsymbol{P}(t;0,\ \boldsymbol{P}_0)=\boldsymbol{P}_+ \quad {}^\forall \boldsymbol{P}_0>\boldsymbol{P}_-$ (66)

ii) $\lim\limits_{t\to+\infty} \boldsymbol{P}(t;0,\ \boldsymbol{P}_0)=\boldsymbol{P}_- \quad {}^\forall \boldsymbol{P}_0<\boldsymbol{P}_+$ (67)

(证明) 首先注意到，在 $\boldsymbol{P}_+>\boldsymbol{P}_-$ 成立的情况下，由引理 7 i)，$\mathrm{Re}\,\lambda(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}_-)>0$ 成立.

若(21)式的解 $\boldsymbol{P}(t;0,\ \boldsymbol{P}_0)$ 存在，由(37)式可见，$\Delta\boldsymbol{P}\triangleq\boldsymbol{P}(t;0,\ \boldsymbol{P}_0)-\boldsymbol{P}_-$ 满足微分方程式

$$-\frac{d}{dt}\Delta\boldsymbol{P}=\Delta\boldsymbol{P}\boldsymbol{A}_-+(\boldsymbol{A}_-)^T\Delta\boldsymbol{P}-\Delta\boldsymbol{P}\boldsymbol{B}\boldsymbol{B}^T\Delta\boldsymbol{P} \tag{68a}$$

$$\Delta\boldsymbol{P}(0)=\boldsymbol{P}_0-\boldsymbol{P}_->\boldsymbol{0} \quad (\boldsymbol{A}_-\triangleq\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}_-) \tag{68b}$$

反之也容易看到，若(68)式的解 $\Delta\boldsymbol{P}$ 存在，则(21)式的解存在，且可以表示成

$$\boldsymbol{P}(t;0,\ \boldsymbol{P}_0)=\boldsymbol{P}_-+\Delta\boldsymbol{P}. \tag{69}$$

但是根据定理 3，(68)式的解 $\Delta\boldsymbol{P}(t)$ 对于任意 $\boldsymbol{P}_0>\boldsymbol{P}_-$，在 $-\infty<t\leqslant0$ 存在，而且是正定的. 因此，对(68a)式左、右乘以 $(\Delta\boldsymbol{P})^{-1}$，若令 $\boldsymbol{\Sigma}\triangleq(\Delta\boldsymbol{P})^{-1}$，可见 $\boldsymbol{\Sigma}$ 满足

$$\dot{\boldsymbol{\Sigma}}=\boldsymbol{\Sigma}(\boldsymbol{A}_-)^T+\boldsymbol{A}_-\boldsymbol{\Sigma}-\boldsymbol{B}\boldsymbol{B}^T, \tag{70}$$

$$\boldsymbol{\Sigma}(0)=(\boldsymbol{P}_0-\boldsymbol{P}_-)^{-1} \quad {}^\forall\boldsymbol{P}_0>\boldsymbol{P}_-. \tag{71}$$

在这里注意到 $\mathrm{Re}\,\lambda(\boldsymbol{A}_-)>0$，利用系 1 的(3)式及引理 7 iii)，则

$$\lim_{t\to-\infty}\boldsymbol{\Sigma}(t)=\int_{-\infty}^{0}e^{(A_-)^Tt}\boldsymbol{B}\boldsymbol{B}^Te^{A_-t}\,dt=(\boldsymbol{P}_+-\boldsymbol{P}_-)^{-1} \quad {}^\forall\boldsymbol{P}_0>\boldsymbol{P}_- \tag{72}$$

成立. 由(69)，(72)式得

$$\lim_{t\to-\infty}\boldsymbol{P}(t;0,\ \boldsymbol{P}_0)=\boldsymbol{P}_-+\lim_{t\to-\infty}\boldsymbol{\Sigma}(t)^{-1}=\boldsymbol{P}_-+(\lim_{t\to-\infty}\boldsymbol{\Sigma}(t))^{-1}=\boldsymbol{P}_+ \quad {}^\forall\boldsymbol{P}_0>\boldsymbol{P}_- \tag{73}$$

证明完毕

ii) 对应于(37)，(21)式，考虑代数方程式

$$-\tilde{\boldsymbol{P}}\boldsymbol{A}-\boldsymbol{A}^T\tilde{\boldsymbol{P}}-\tilde{\boldsymbol{P}}\boldsymbol{B}\boldsymbol{B}\tilde{\boldsymbol{P}}+\boldsymbol{Q}=\boldsymbol{0} \tag{74}$$

及微分方程式

$$-\dot{\tilde{\boldsymbol{P}}}=-\tilde{\boldsymbol{P}}\boldsymbol{A}-\boldsymbol{A}^T\tilde{\boldsymbol{P}}-\tilde{\boldsymbol{P}}\boldsymbol{B}\boldsymbol{B}^T\tilde{\boldsymbol{P}}+\boldsymbol{Q}$$

$$\tilde{\boldsymbol{P}}(0)=\tilde{\boldsymbol{P}}_0. \tag{75}$$

显然，若(37)式具有实对称解，则(74)式也具有实对称解. 而且，若令其最大(小)解为 $\tilde{\boldsymbol{P}}_+(\tilde{\boldsymbol{P}}_-)$，则

$$\tilde{\boldsymbol{P}}_+=-\boldsymbol{P}_-,\ \tilde{\boldsymbol{P}}_-=-\boldsymbol{P}_+. \tag{76}$$

因此，根据 i)，对于(75)式的解 $\tilde{\boldsymbol{P}}(t;0,\ \tilde{\boldsymbol{P}}_0)$，

$$\lim_{t\to-\infty}\tilde{\boldsymbol{P}}(t;0,\ \boldsymbol{P}_0)=\tilde{\boldsymbol{P}}_+ \quad {}^\forall\tilde{\boldsymbol{P}}_0>\tilde{\boldsymbol{P}}_- \tag{77}$$

成立. 在这里若令 $\boldsymbol{P}(t)\triangleq-\tilde{\boldsymbol{P}}(-t;0,\ \tilde{\boldsymbol{P}}_0)$，首先根据(75)式，$\boldsymbol{P}$ 是微分方程式

$$-\dot{\boldsymbol{P}}=\boldsymbol{P}\boldsymbol{A}+\boldsymbol{A}^T\boldsymbol{P}-\boldsymbol{P}\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}+\boldsymbol{Q}$$

$$\boldsymbol{P}(0)=-\tilde{\boldsymbol{P}}_0 \tag{78}$$

的解，而且根据(76)，(77)式，

$$\lim_{t\to-\infty}\boldsymbol{P}(t)=\boldsymbol{P}_- \quad {}^\forall\boldsymbol{P}(0)<\boldsymbol{P}_+$$

成立.

证明完毕

以该结果为基础，容易证明(参看附录 B)在 $\boldsymbol{P}_+>\boldsymbol{P}_-$ 成立的情况下，$\boldsymbol{P}_+(\boldsymbol{P}_-)$ 在负(正)的时间方向局部渐近稳定，即对于足够小的正数 ε,

$$\|\boldsymbol{P}_0-\boldsymbol{P}_+\|<\varepsilon\Rightarrow\lim_{t\to-\infty}\boldsymbol{P}(t;0,\boldsymbol{P}_0)=\boldsymbol{P}_+ \tag{79a}$$

$$(\|\boldsymbol{P}_0-\boldsymbol{P}_-\|<\varepsilon\Rightarrow\lim_{t\to\infty}\boldsymbol{P}(t;0,\boldsymbol{P}_0)=\boldsymbol{P}_- \tag{79b}$$

成立. (参照附录 B)

而且，关于 P_+, P_- 以外的平衡点 $\boldsymbol{P}(\boldsymbol{P}_-\leqslant\boldsymbol{P}\leqslant\boldsymbol{P}_+)$，例如下列 $\boldsymbol{P}$ 邻近的点，因

$$\boldsymbol{P}_0\triangleq\boldsymbol{P}+\varepsilon\boldsymbol{I}\quad(\boldsymbol{P}_0\triangleq\boldsymbol{P}-\varepsilon\boldsymbol{I})\,(0<\varepsilon<1)$$

满足

$$\boldsymbol{P}_0\geqslant\boldsymbol{P}_-+\varepsilon\boldsymbol{I}>\boldsymbol{P}_-\quad(\boldsymbol{P}_0\leqslant\boldsymbol{P}_+-\varepsilon\boldsymbol{I}<\boldsymbol{P}_+)$$

及定理 8, 可见 $\boldsymbol{P}$ 不是稳定的平衡点.

(注意)

i) 如上所述，可以证明微分方程式(21)的解对于满足

$$\boldsymbol{P}_0>\boldsymbol{P}_-\quad(\boldsymbol{P}_0<\boldsymbol{P}_+) \tag{80}$$

的起始值 $\boldsymbol{P}_0$, 渐近于局部渐近稳定的平衡点 $\boldsymbol{P}_+(\boldsymbol{P}_-)$. 也可以证明[1), 实际上其逆亦成立. 即，若 $\boldsymbol{P}_+>\boldsymbol{P}_-$, 则 $\lim_{t\to-\infty}\boldsymbol{P}(t;0,\boldsymbol{P}_0)=\boldsymbol{P}_+\Rightarrow\boldsymbol{P}_0>\boldsymbol{P}_-(\lim_{t\to\infty}\boldsymbol{P}(t;0,\boldsymbol{P}_0)=\boldsymbol{P}_-\Rightarrow\boldsymbol{P}_0<\boldsymbol{P}_+)$ 成立.

ii) 若可以找到一个满足(80)式的 $\boldsymbol{P}_0$, 则由(66), (67)式可以求出代数方程式(37)的最大(小)解 $\boldsymbol{P}_+(\boldsymbol{P}_-)$. 因 $\boldsymbol{P}_+$, $\boldsymbol{P}_-$ 当然是未知的，所以找出这样的 $\boldsymbol{P}_0$ 一般不是简单的. 下面关于其求法再稍谈一下. 若预先保证 $\boldsymbol{P}_+>\boldsymbol{0}>\boldsymbol{P}_-$, 当然选取 $\boldsymbol{P}_0=\boldsymbol{0}$ 即可. 例如，它相当于代数方程式(41)中$(\boldsymbol{A},\boldsymbol{B})$可控且$(\boldsymbol{C},\boldsymbol{A})$可观测的情况(其证明参照附录 C). 对于一般的方程式(37), 例如，若在(58)式的解中设满足(59)式的解为 $\boldsymbol{P}_0$, 首先根据定理 5, 因 $\boldsymbol{P}_0$ 本身就是(58)式的最大解，而且是正定的，根据前述，则 $\boldsymbol{P}_0$ 可以作为某个微分方程式的极限求出. 其次考虑到(57)式，若利用引理 7 的 iv), 则 $\boldsymbol{P}_0>\boldsymbol{P}_-$ 成立. 因而 $\boldsymbol{P}_0$ 是所要求的. 同样，在(58)式的解 $\boldsymbol{P}$ 中，若令满足 $\mathrm{Re}\,\lambda(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P})>0$ 的解为 $\boldsymbol{P}_0$, 则 $\boldsymbol{P}_0<\boldsymbol{P}_+$.

iii) 如注意 ii) 中所述，若$(\boldsymbol{A},\boldsymbol{B})$可控，$(\boldsymbol{C},\boldsymbol{A})$可观测，则对应于代数方程式(41)的微分方程式(26)的解 $\boldsymbol{P}(t;0,\boldsymbol{P}_0)$ 满足

$$\lim_{t\to-\infty}\boldsymbol{P}(t;0,\boldsymbol{P}_0)=\boldsymbol{P}_+\quad{}^\forall\boldsymbol{P}_0\geqslant 0. \tag{81}$$

实际上，在更宽的$(\boldsymbol{A},\boldsymbol{B})$可稳定、$(\boldsymbol{C},\boldsymbol{A})$可检出的假定下，可以证明该性质亦成立(证明参照附录 D). 注意，最后若将该结果、定理 5 及引理 5 的(ii), (v)合起来，可以证明其结果是以前讲过的常系数最优调节器问题有关性质的一部分(P-I-6 中定理 3(b), (c), (f) 及(d), (e)的充分性).

---

1) 首先，设(37)式的任意实对称解为 $\boldsymbol{P}$ 时

i) $\boldsymbol{P}_0\geqslant(>)\boldsymbol{P}\Rightarrow\boldsymbol{P}(t;0,\boldsymbol{P}_0)\geqslant(>)\boldsymbol{P}\quad{}^\forall t$

ii) $\boldsymbol{P}_0\leqslant(<)\boldsymbol{P}\Rightarrow\boldsymbol{P}(t;0,\boldsymbol{P}_0)\leqslant(<)\boldsymbol{P}\quad{}^\forall t$

iii) $(\boldsymbol{P}_0-\boldsymbol{P})$不是定值$\Rightarrow\boldsymbol{P}(t;0,\boldsymbol{P}_0)-\boldsymbol{P}$ 对于任意 t 也不是定值的性质成立. 关于这一点，若在(68)式中令 $\Delta\boldsymbol{P}=\boldsymbol{P}(t;0,\boldsymbol{P}_0)-\boldsymbol{P}$, 对该式利用系 1, 则可以很容易证明. 根据该性质和 $\boldsymbol{P}_+>\boldsymbol{P}_-$ 的假定，其逆也很容易证明.

2.4 离散型黎卡提方程式

对于2.1节中给出的矩阵 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{Q}$(但设 $\boldsymbol{Q}\geqslant\boldsymbol{0}$)，若考虑到下列非线性代数方程式[1]

$$\boldsymbol{P}=\boldsymbol{A}^T\boldsymbol{P}\boldsymbol{A}-\boldsymbol{A}^T\boldsymbol{P}\boldsymbol{B}(\boldsymbol{I}+\boldsymbol{B}^T\boldsymbol{P}\boldsymbol{B})^{-1}\boldsymbol{B}^T\boldsymbol{P}\boldsymbol{A}+\boldsymbol{Q}, \tag{82}$$

则利用某种变换可以将该方程式归结到2.2节中讲过的黎卡提代数方程式(37)形式求解．下面我们来证明．

首先，对于给定的 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{Q}$，考虑下列变换

$$\boldsymbol{F}=(\boldsymbol{A}+\boldsymbol{I})^{-1}(\boldsymbol{A}-\boldsymbol{I})$$

$$\boldsymbol{G}=2(\boldsymbol{A}+\boldsymbol{I})^{-2}\boldsymbol{B}$$

$$\boldsymbol{W}_{11}=\boldsymbol{Q}$$

$$\boldsymbol{W}_{12}=-\boldsymbol{Q}(\boldsymbol{A}+\boldsymbol{I})^{-1}\boldsymbol{B}$$

$$\boldsymbol{W}_{22}=\boldsymbol{I}+\boldsymbol{B}^T(\boldsymbol{A}^T+\boldsymbol{I})^{-1}\boldsymbol{Q}(\boldsymbol{A}+\boldsymbol{I})^{-1}\boldsymbol{B}(>\boldsymbol{0}),$$

其次，对于由该变换得到的矩阵 $\boldsymbol{F}, \boldsymbol{G}, \boldsymbol{W}_{11}, \boldsymbol{W}_{12}, \boldsymbol{W}_{22}$，考虑代数方程式

$$\hat{\boldsymbol{P}}\boldsymbol{F}+\boldsymbol{F}^T\hat{\boldsymbol{P}}-(\hat{\boldsymbol{P}}\boldsymbol{G}+\boldsymbol{W}_{12})\boldsymbol{W}_{22}^{-1}(\hat{\boldsymbol{P}}\boldsymbol{G}+\boldsymbol{W}_{12})^T+\boldsymbol{W}_{11}=\boldsymbol{0},$$

该方程式可以很容易地改写成(37)式的形式

$$\hat{\boldsymbol{P}}\hat{\boldsymbol{A}}+\hat{\boldsymbol{A}}^T\hat{\boldsymbol{P}}-\hat{\boldsymbol{P}}\hat{\boldsymbol{B}}\hat{\boldsymbol{B}}^T\hat{\boldsymbol{P}}+\hat{\boldsymbol{Q}}=\boldsymbol{0} \tag{83}$$

式中 $\hat{\boldsymbol{A}}\triangleq\boldsymbol{F}-\boldsymbol{G}\boldsymbol{W}_{22}^{-1}\boldsymbol{W}_{12}$

$\hat{\boldsymbol{B}}\triangleq\boldsymbol{G}\boldsymbol{W}_{22}^{-1/2}$

$\hat{\boldsymbol{Q}}\triangleq\boldsymbol{W}_{11}-\boldsymbol{W}_{12}\boldsymbol{W}_{22}^{-1}\boldsymbol{W}_{22}^T(\geqslant\boldsymbol{Q})$[2]

那么，可以证明下列性质(试证明)

i) (75)式的实对称解 $\hat{\boldsymbol{P}}$ 存在，且

ii) 令 $\boldsymbol{P}\triangleq\frac{1}{2}(\boldsymbol{A}^T+\boldsymbol{I})^{-1}\hat{\boldsymbol{P}}(\boldsymbol{A}+\boldsymbol{I})^{-1}$ 时

满足 $\boldsymbol{I}+\boldsymbol{B}^T\boldsymbol{P}\boldsymbol{B}>\boldsymbol{0}\Rightarrow\boldsymbol{P}$ 是(82)式的实对称解

(注意)

若$(\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}})$可稳定，根据定理5，因(83)式的准正定实对称解 $\hat{\boldsymbol{P}}$ 存在，故满足i)，ii)的条件．

其3 黎卡提非线性方程式的普遍形式

在前一节考察了标准黎卡提型代数方程式(37)式，下面再考虑一下更普遍的形式．

首先，对应于代数方程式(37)，考虑代数方程式

$$\boldsymbol{P}\boldsymbol{A}+\boldsymbol{A}^T\boldsymbol{P}-(\boldsymbol{P}\boldsymbol{B}+\boldsymbol{W}_{12})\boldsymbol{W}_{22}^{-1}(\boldsymbol{P}\boldsymbol{B}+\boldsymbol{W}_{12})^T+\boldsymbol{W}_{11}=\boldsymbol{0} \tag{84}$$

其中假定，$\boldsymbol{A}\in R^{n\times n}$，$\boldsymbol{B}\in R^{n\times m}$，而 $\boldsymbol{W}_{11}\in R^{n\times n}$，$\boldsymbol{W}_{22}\in R^{m\times m}$ 均为实对称矩阵，且 $\boldsymbol{W}_{22}>\boldsymbol{0}$．显然，(37)式相当于在(84)式中特别限定 $\boldsymbol{W}_{12}=\boldsymbol{0}$，$\boldsymbol{W}_{22}=\boldsymbol{I}$，将 $\boldsymbol{W}_{11}$ 写成 $\boldsymbol{Q}$ 的情况．

---

1) 以前讲过的黎卡提方程式(37)是对应于连续系统的问题(调节器问题，滤波器问题等)，而该方程式与离散系统的问题有关，故称为离散型黎卡提方程式．

2) 根据变换的定义

$$\begin{bmatrix}\boldsymbol{W}_{11} & \boldsymbol{W}_{12}\\ \boldsymbol{W}_{12}^T & \boldsymbol{W}_{22}\end{bmatrix}=\begin{bmatrix}\boldsymbol{I} & -(\boldsymbol{A}+\boldsymbol{I})^{-1}\boldsymbol{B}\\ \boldsymbol{0} & \boldsymbol{I}\end{bmatrix}^T\begin{bmatrix}\boldsymbol{Q} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}\end{bmatrix}\begin{bmatrix}\boldsymbol{I} & -(\boldsymbol{A}+\boldsymbol{I})^{-1}\boldsymbol{B}\\ \boldsymbol{0} & \boldsymbol{I}\end{bmatrix}\geqslant\boldsymbol{0}$$

$\Leftrightarrow\hat{\boldsymbol{Q}}\geqslant\boldsymbol{0}$(参照B-I-14中性质8a)

为了方便起见，定义如下合成矩阵 $\boldsymbol{W}$

$$\boldsymbol{W}\triangleq\begin{array}{cc} \begin{array}{cc} n & m \end{array} & \\ \left[\begin{array}{c|c} \widehat{\boldsymbol{W}_{11}} & \widehat{\boldsymbol{W}_{12}} \\ \hline \boldsymbol{W}_{12}^T & \boldsymbol{W}_{22} \end{array}\right] & \begin{array}{l})\,n \\ )\,m \end{array} \end{array}=\boldsymbol{W}^T\in R^{(n+m)\times(n\times m)}.$$

其次，对于前述矩阵 $\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{W}$，设 $\boldsymbol{P}=\boldsymbol{P}^T\in R^{n\times n}$ 为未知矩阵，考虑非线性矩阵不等式

$$\boldsymbol{PA}+\boldsymbol{A}^T\boldsymbol{P}-(\boldsymbol{PB}+\boldsymbol{W}_{12})\boldsymbol{W}_{22}^{-1}(\boldsymbol{PB}+\boldsymbol{W}_{12})^T+\boldsymbol{W}_{11}\geqslant\boldsymbol{0} \tag{85}$$

在(84)式的解必定是(85)式的解的意义上，(85)式是(84)式更普遍的形式.

但是，若将由给定的$(\boldsymbol{A},\ \boldsymbol{B};\ \boldsymbol{W})$和实对称矩阵 $\boldsymbol{P}$ 构成的合成矩阵定义成

$$\boldsymbol{F}(\boldsymbol{P};\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{W})\triangleq\begin{bmatrix}\boldsymbol{PA}+\boldsymbol{A}^T\boldsymbol{P}+\boldsymbol{W}_{11} & \boldsymbol{PB}+\boldsymbol{W}_{12}\\ \boldsymbol{B}^T\boldsymbol{P}+\boldsymbol{W}_{12}^T & \boldsymbol{W}_{22}\end{bmatrix},$$

则 $\boldsymbol{P}$ 是(85)式的实对称解和对于给定的$(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{W})$，$\boldsymbol{P}$ 是满足线性矩阵不等式

$$\boldsymbol{F}(\boldsymbol{P};\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{W})\geqslant\boldsymbol{0} \tag{86}$$

的实对称解等价.

（问题 2） 试证明上述结论（提示：利用 B-I-14 中性质 8a，并且注意到，因 $\boldsymbol{W}_{22}>\boldsymbol{0}$，则 $\boldsymbol{W}_{22}$ 的逆矩阵存在）.

(84)式的解都是(85)式的解，因而也是(86)式的解. 但是反过来，(86)式的解却不一定都是(84)式的解[1)]. 在(86)式的解中，仅 rank $\boldsymbol{F}(\boldsymbol{P};\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{W})$ 等于 rank $\boldsymbol{W}_{22}$，即取最小值m的才是(84)式的解.

（问题 3） 试证明上述结论（提示：利用 B-I-14 中性质 8b，而且注意到，因 $\boldsymbol{W}_{22}>\boldsymbol{0}$，则对于任意的 $\boldsymbol{P}\in R^{n\times n}$，rank $\boldsymbol{F}(\boldsymbol{P};\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{W})\geqslant$rank $\boldsymbol{W}_{22}=m$ 成立）.

(86)式对于未知矩阵 $\boldsymbol{P}$ 是线性的，因而它比(85)式易于处理，所以下面我们来考察(86)式实对称解的存在及其性质[2)].

代数方程式(37)实对称解的存在条件是不等式(61)成立（参看定理 7）. 因此，与(37)式的普遍式(84)相对应，我们来考虑(61)式的普遍式

$$\begin{aligned}\boldsymbol{H}(\omega;\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{W})\triangleq\ &\boldsymbol{W}_{22}+\boldsymbol{B}^T(-i\omega\boldsymbol{I}-\boldsymbol{A}^T)^{-1}\boldsymbol{W}_{12}+\boldsymbol{W}_{12}^T(i\omega\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}\\ &+\boldsymbol{B}^T(-i\omega\boldsymbol{I}-\boldsymbol{A}^T)^{-1}\boldsymbol{W}_{11}(i\omega\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}\geqslant\boldsymbol{0}\quad \text{a. e. } \omega\in R\end{aligned} \tag{87}$$

实际上，(87)式是(84)式实对称解存在的条件. 为了推导出该结果，首先对于给定的矩阵组$(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{W})$考虑表 1 所示的四种变换.

而且，若将这四种变换按着(iv)，(iii)，(ii)，(i)的顺序进行表示成$(\boldsymbol{T},\ \boldsymbol{G},\ \boldsymbol{K},\ \boldsymbol{P}_0)$，则容易证明由变换$(\boldsymbol{T},\ \boldsymbol{G},\ \boldsymbol{K},\ \boldsymbol{P}_0)$可将$(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{W})$变换成

$$(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{W})\xrightarrow{(\boldsymbol{T},\ \boldsymbol{G},\ \boldsymbol{K},\ \boldsymbol{P}_0)}(\tilde{\boldsymbol{A}},\ \tilde{\boldsymbol{B}},\ \tilde{\boldsymbol{W}})$$

式中 $\tilde{\boldsymbol{A}}=\boldsymbol{T}^{-1}\boldsymbol{A}_k\boldsymbol{T}\quad(\boldsymbol{A}_k\triangleq\boldsymbol{A}+\boldsymbol{BK})$ (88a)

---

1) 例如，若设 $n=m=1$，$\boldsymbol{W}_{11}=\boldsymbol{W}_{22}=1$，$\boldsymbol{W}_{12}=0$，则 $\boldsymbol{P}=0$ 满足(86)式，但不满足(84)式.

2) 因这里仍限定 $\boldsymbol{W}_{22}>\boldsymbol{0}$，则(85)式和(86)式等价. 但是若将对 $\boldsymbol{W}_{22}$ 的限制放宽到 $\boldsymbol{W}_{22}\geqslant\boldsymbol{0}$，则(86)式的解是在(86)式中将 $\boldsymbol{W}_{22}^{-1}$ 用其广义逆矩阵 $\boldsymbol{W}_{22}$ 置换后的解，但其逆不一定成立（参看 B-I-14 中性质 8a）. 亦即，若包含到 $\boldsymbol{W}_{22}\geqslant\boldsymbol{0}$ 的情况，从普遍的程度来看，则可以认为(86)式是处于(84)式和(85)式中间.

表1 变换$(A, B, W) \to (\tilde{A}, \tilde{B}, \tilde{W})$的定义

($T \in R^{n\times n}$, $G \in R^{m\times m}$, $K \in R^{m\times n}$, $P_0 = P_0^T \in R^{n\times n}$, 而且 T 和 G 是正则的)

变换后的矩阵	变			换
	(i)	(ii)	(iii)	(iv)
	状态的座标变换	输入的座标变换	反 馈 变 换	P—变换[154]
$\tilde{A}$	$T^{-1}AT$	A	$A+BK$	A
$\tilde{B}$	$T^{-1}B$	BG	B	B
$\tilde{W}_{11}$	$T^T W_{11} T$	W_{11}	$W_{11}+K^T W_{12}^T + W_{12}K + K^T W_{22} K$	$P_0 A + A^T P_0 + W_{11}$
$\tilde{W}_{12}$	$T^T W_{12}$	$W_{12}G$	$W_{12}+K^T W_{22}$	$P_0 B + W_{12}$
$\tilde{W}_{22}$	W_{22}	$G^T W_{22} G$	W_{22}	W_{22}

$$\tilde{B} = T^{-1}BG \tag{88b}$$

$$\tilde{W} \triangleq \begin{bmatrix} \tilde{W}_{11} & \tilde{W}_{12} \\ \tilde{W}_{12}^T & \tilde{W}_{22} \end{bmatrix} = \begin{bmatrix} T & 0 \\ KT & G \end{bmatrix}^T \left[F(P_0; A, B, W) \right] \begin{bmatrix} T & 0 \\ KT & G \end{bmatrix} \tag{88c}$$

在这里，我们将所有这类变换(T, G, K, P_0)(其中假定 $T \in R^{n\times n}$, $K \in R^{m\times n}$, $P_0 = P_0^T \in R^{n\times n}$, 而且 T, G 是正则的)写成 $\mathscr{G}$.

[定义] 对于两个矩阵组(A, B, W)和$(\tilde{A}, \tilde{B}, \tilde{W})$，当存在适当的变换$(T, G, K, P_0) \in \mathscr{G}$, 使得

$$(A, B, W) \xrightarrow{(T, G, K, P_0)} (\tilde{A}, \tilde{B}, \tilde{W})$$

时，称为两个组是等价的，我们将表示成$(A, B, W) \sim (\tilde{A}, \tilde{B}, \tilde{W})$.

(注意)

若将 $\mathscr{G}$ 的任意两个变换的积定义成

$$(\tilde{T}, \tilde{G}, \tilde{K}, \tilde{P}_0) \circ (T, G, K, P_0) = (T\tilde{T}, G\tilde{G}, K + G\tilde{K}T^{-1}, P_0 + (T^{-1})^T \tilde{P}_0 T^{-1}),$$

则 $\mathscr{G}$ 是关于"$\circ$"的群. 而且，"$\sim$"是等值关系(试证明).

关于相互等价的两个矩阵组有如下性质.

[引理10] 设$(A, B, W) \sim (\tilde{A}, \tilde{B}, \tilde{W})$，即假定存在适当的$(T, G, K, P_0)$，使得(88)式成立.

i) 对于某个 $P = P^T \in R^{n\times n}$. 若 $F(P; A, B, W) \geq 0$ 成立，设 $\tilde{P} \triangleq T^T(P - P_0)T$, 则 $F(\tilde{P}; \tilde{A}, \tilde{B}, \tilde{W}) \geq 0$ 成立.

ii) 对于某个 $P = P^T \in R^{n\times n}$, 若(84)式成立，设 $\tilde{P} \triangleq T^T(P - P_0)T$, 则下式成立

$$\tilde{P}\tilde{A} + \tilde{A}^T\tilde{P} - (\tilde{P}\tilde{B} + \tilde{W}_{12})\tilde{W}_{22}^{-1}(\tilde{P}\tilde{B} + \tilde{W}_{12})^T + \tilde{W}_{11} = 0 \tag{89}$$

iii) $H(\omega; A, B, W) \geq 0$ (a. e. $\omega \in R$)

$\Leftrightarrow H(\omega; \tilde{A}, \tilde{B}, \tilde{W}) \geq 0$ (a. e. $\omega \in R$)

(证明) i) 首先，根据 P, P_0 的实对称性和 $\tilde{P}$ 的定义，显然 $\tilde{P}$ 是实对称的. 而且，根据(88)式和 $\tilde{P}$ 的定义，容易证明

$$F(\widetilde{P}, \widetilde{A}, \widetilde{B}, \widetilde{W})=\begin{bmatrix} T & 0 \\ KT & G \end{bmatrix}^T [F(P; A, B, W)]\begin{bmatrix} T & 0 \\ KT & G \end{bmatrix} \tag{90}$$

因此，下列同值关系成立

$$F(P; A, B, W)\geqslant 0 \Leftrightarrow F(\widetilde{P}; \widetilde{A}, \widetilde{B}, \widetilde{W})\geqslant 0$$

ii) 将 $\widetilde{P}, \widetilde{A}, \widetilde{B}, \widetilde{W}$ 代入(89)式左边，由(84)式易见(89)式成立.

iii) 首先注意，利用任意的 $P_0=P_0^T\in R^{n\times n}$, (87)式左边可以表示成

$$H(\omega; A, B, W)=\begin{bmatrix} (-i\omega I-A)^{-1}B \\ I \end{bmatrix}^T [F(P_0; A, B, W)]\begin{bmatrix} (i\omega I-A)^{-1}B \\ I \end{bmatrix} \tag{91}$$

在这里若注意到 $F(0; A, B, W)=W$ 则可得

$$H(\omega; \widetilde{A}, \widetilde{B}, \widetilde{W})=\begin{bmatrix} (-i\omega I-\widetilde{A})^{-1}\widetilde{B} \\ I \end{bmatrix}^T [\widetilde{W}]\begin{bmatrix} (i\omega I-\widetilde{A})^{-1}\widetilde{B} \\ I \end{bmatrix}$$

将(88c)代入该式，得

$$H(\omega; \widetilde{A}, \widetilde{B}, \widetilde{W})=\begin{bmatrix} C(-i\omega) \\ D(-i\omega) \end{bmatrix}^T [F(P_0; A, B, W)]\begin{bmatrix} C(i\omega) \\ D(i\omega) \end{bmatrix}$$

式中

$$C(i\omega)\triangleq T(i\omega I-\widetilde{A})^{-1}\widetilde{B}$$
$$D(i\omega)\triangleq KT(i\omega I-\widetilde{A})^{-1}\widetilde{B}+G$$

在这里根据(88a)，(88b)式[1]，因为 $D(i\omega)$, $C(i\omega)$ 可表示成

$$\begin{aligned} D(i\omega)&=K(i\omega I-A_k)^{-1}BG+G \\ C(i\omega)&=[(i\omega I-A)-BK]^{-1}BG \\ &=(i\omega I-A)^{-1}[I-BK(i\omega I-A)^{-1}]^{-1}BG \\ &=(i\omega I-A)^{-1}[I+BK(i\omega I-A_k)^{-1}]BG \\ &=(i\omega I-A)^{-1}BD(i\omega) \end{aligned}$$

将其代入上式，得

$$\begin{aligned} &\widetilde{H}(\omega; \widetilde{A}, \widetilde{B}, \widetilde{W}) \\ &=D(-i\omega)^T\begin{bmatrix} (-i\omega I-A)^{-1}B \\ I \end{bmatrix}^T [F(P_0; A, B, W)]\begin{bmatrix} (i\omega I-A)^{-1}B \\ I \end{bmatrix}D(i\omega). \end{aligned}$$

利用上面讲过的(91)式，则得

$$\widetilde{H}(\omega; \widetilde{A}, \widetilde{B}, \widetilde{W})=D(-i\omega)^T H(\omega; A, B, W)D(i\omega) \quad (\text{a. e. } \omega\in R)$$

因此

$$\begin{aligned} &H(\omega; A, B, W)\geqslant 0 \quad (\text{a. e. } \omega\in R) \\ \Leftrightarrow &H(\omega; \widetilde{A}, \widetilde{B}, \widetilde{W})\geqslant 0 \quad (\text{a. e. } \omega\in R) \end{aligned}$$

证明完毕

由该结果可见，线性矩阵不等式(86)及代数方程式(84)实对称解的存在性，对于前述变换不变. 因此，在讨论(84)，(86)式实对称解的存在及求其解的情况下，可以利用适当变换将这些式子变换成易于处理的形式. 例如

1) 参照 B-I-1 中定理 3 的性质(vi)中讲过的逆矩阵的公式.

a) 由变换(i)，变换(ii)，(i)，变换(iii)，(ii)，(i)等，将$(\boldsymbol{A}, \boldsymbol{B})$变成典范形(参照 A-I-12, A-I-13)；

b) 由变换$(\boldsymbol{I}, \boldsymbol{W}_{22}^{-1/2}, \boldsymbol{0}, \boldsymbol{0})$，使$\widetilde{\boldsymbol{W}}_{22}=\boldsymbol{I}$;

c) 由变换$(\boldsymbol{I}, \boldsymbol{I}, -\boldsymbol{W}_{22}^{-1}\boldsymbol{W}_{12}^{T}, \boldsymbol{0})$，使$\widetilde{\boldsymbol{W}}_{12}=\boldsymbol{0}$;

d) 若 B-I-15 中其 1 讲过的线性方程式$\boldsymbol{X}\boldsymbol{A}+\boldsymbol{A}^{T}\boldsymbol{X}=-\boldsymbol{W}_{11}$具有实对称解$\boldsymbol{X}$,则将$\boldsymbol{X}$选为变换(iv)中的$\boldsymbol{P}_0$,使$\widetilde{\boldsymbol{W}}_{11}\triangleq\boldsymbol{P}_0\boldsymbol{A}+\boldsymbol{A}^{T}\boldsymbol{P}+\boldsymbol{W}_{11}=\boldsymbol{0}$;

e) 若线性不等式(86)具有实对称解,则将它选为变换(iv)中的$\boldsymbol{P}_0$,使

$$\widetilde{\boldsymbol{W}}\triangleq\boldsymbol{F}(\boldsymbol{P}_0; \boldsymbol{A}, \boldsymbol{B}, \boldsymbol{W})\geqslant 0,$$

等等,均可实现.

此外可以看到,若将 b),c) 合起来,则这里考虑的矩阵组$(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{W})$由变换$(\boldsymbol{I}, \boldsymbol{W}_{22}^{-1/2}, -\boldsymbol{W}_{22}^{-1}\boldsymbol{W}_{12}^{T}, 0)$常可变换成在其 2 中讨论过的(37)式的形式(为了方便起见，称其为“准标准形”,该变换称为“准标准化变换”)

$$(\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{B}}, \widetilde{\boldsymbol{Q}}\dotplus\boldsymbol{I}), \widetilde{\boldsymbol{Q}}=\widetilde{\boldsymbol{Q}}^{T} \quad (\widetilde{\boldsymbol{Q}}\triangleq\boldsymbol{W}_{11}-\boldsymbol{W}_{12}\boldsymbol{W}_{22}^{-1}\boldsymbol{W}_{12}^{T}).$$

在准标准形中，一般$\widetilde{\boldsymbol{Q}}\geqslant\boldsymbol{0}$不一定成立，而在以前的$(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{W})$中，若$\boldsymbol{W}\geqslant\boldsymbol{0}$，则在由此得到的准标准形中，$\widetilde{\boldsymbol{Q}}\triangleq\boldsymbol{W}_{11}-\boldsymbol{W}_{12}\boldsymbol{W}_{22}^{-1}\boldsymbol{W}_{12}^{T}\geqslant\boldsymbol{0}$(参照 B-I-14 中广义逆矩阵的性质 8b).由此可见,在(86)式具有实对称解的情况下，首先利用 e)中讲过的变换(iv)将$(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{W})$变换成$(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{F}(\boldsymbol{P}_0; \boldsymbol{A}, \boldsymbol{B}, \boldsymbol{W}))$(显然，$\boldsymbol{F}(\boldsymbol{P}_0; \boldsymbol{A}, \boldsymbol{B}, \boldsymbol{W})\geqslant\boldsymbol{0}$)，然后再利用准标准化变换将其变成(41)式最标准的形式(此形称为“标准形”,此变换称为“标准化变换”).

$$(\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{B}}, \widetilde{\boldsymbol{Q}}\dotplus\boldsymbol{I}), \widetilde{\boldsymbol{Q}}=\widetilde{\boldsymbol{Q}}^{T}\geqslant\boldsymbol{0}.$$

反之,若$(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{W})$可用属于$\mathscr{G}$的变换变成标准形$(\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{B}}, \widetilde{\boldsymbol{W}})$,则 i) 由其逆变换可使$(\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{B}}, \widetilde{\boldsymbol{W}})$变为$(\boldsymbol{A},\boldsymbol{B}, \boldsymbol{W})$; ii) 关于标准形$(\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{B}}, \widetilde{\boldsymbol{W}})$,常存在满足$\boldsymbol{F}(\boldsymbol{P}; \widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{B}}, \widetilde{\boldsymbol{W}})\geqslant\boldsymbol{0}$的实对称解(例如$\boldsymbol{P}=\boldsymbol{0}$),由该性质和引理 10 i) 可见，$\boldsymbol{F}(\boldsymbol{P}; \boldsymbol{A}, \boldsymbol{B}, \boldsymbol{W})\geqslant\boldsymbol{0}$也具有实对称解.

归纳以上,得

[系 3]　给定的$(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{W})$与标准形等价的充分必要条件是，线性矩阵不等式(86)具有实对称解[1].

但是,(84)式可用以前讲过的标准化变换变成(37)式形式的代数方程式

$$\boldsymbol{P}\widetilde{\boldsymbol{A}}+\widetilde{\boldsymbol{A}}^{T}\boldsymbol{P}-\boldsymbol{P}\widetilde{\boldsymbol{B}}\widetilde{\boldsymbol{B}}^{T}\boldsymbol{P}+\widetilde{\boldsymbol{Q}}=\boldsymbol{0} \tag{92}$$

式中　$\widetilde{\boldsymbol{A}}\triangleq\boldsymbol{A}-\boldsymbol{B}\boldsymbol{W}_{22}^{-1}\boldsymbol{W}_{12}^{T}$, $\widetilde{\boldsymbol{B}}\triangleq\widetilde{\boldsymbol{B}}\boldsymbol{W}_{22}^{-1/2}$, $\widetilde{\boldsymbol{Q}}\triangleq\boldsymbol{W}_{11}-\boldsymbol{W}_{12}\boldsymbol{W}_{22}^{-1}\boldsymbol{W}_{12}^{T}$.　因此，若对(90)式利用前节定理 7,则可得如下定理.

[定理 9]　设$(\boldsymbol{A}, \boldsymbol{B})$可控，这时代数方程式(84)的实对称解存在的充分必要条件是(87)式成立．而且,若满足该条件,则

i) 存在满足下列条件的实对称解$\boldsymbol{P}_+$和$\boldsymbol{P}_-$

$$\mathrm{Re}\,\lambda[\boldsymbol{A}-\boldsymbol{B}\boldsymbol{W}_{22}^{-1}(\boldsymbol{B}^{T}\boldsymbol{P}_{+}+\boldsymbol{W}_{12}^{T})]\leqslant\boldsymbol{0}$$

$$\mathrm{Re}\,\lambda[\boldsymbol{A}-\boldsymbol{B}\boldsymbol{W}_{22}^{-1}(\boldsymbol{B}^{T}\boldsymbol{P}_{-}+\boldsymbol{W}_{12}^{T})]\geqslant\boldsymbol{0}$$

---

1) 在$\boldsymbol{W}_{22}\geqslant 0$的情况下,给定的$(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{W})$可由属于$\mathscr{G}$的适当变换变成

$$(\widetilde{\boldsymbol{A}}, \boldsymbol{B}, \widetilde{\boldsymbol{Q}}\dotplus\widetilde{\boldsymbol{R}}), \widetilde{\boldsymbol{Q}}\geqslant\boldsymbol{0}, \widetilde{\boldsymbol{R}}\geqslant\boldsymbol{0}$$

这一点和(86)式具有实对称解等价(若注意到 B-I-14 中性质 8a，证明完全一样,该变换仍用$(\boldsymbol{I}, \boldsymbol{I}, -\boldsymbol{D}-\boldsymbol{C}, \boldsymbol{P}_0)$给出).

ii) $\boldsymbol{P}_+(\boldsymbol{P}_-)$是该方程式的最大解(最小解).

(证明) 首先,若$(\boldsymbol{A}, \boldsymbol{B})$可控,则$(\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{B}})$亦可控(参照 A-I-13 中定理 1). 而且,若$(\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{B}})$可控,根据定理 7,则(92)式,因而(84)式具有实对称解的充分必要条件是

$$\boldsymbol{I}+\tilde{\boldsymbol{B}}^T(-i\omega\boldsymbol{I}-\tilde{\boldsymbol{A}})^{-1}\tilde{\boldsymbol{Q}}(i\omega\boldsymbol{I}-\tilde{\boldsymbol{A}})^{-1}\tilde{\boldsymbol{B}}\geqslant\boldsymbol{0} \quad \text{a. e. } \omega\in R$$

此外,根据引理 10 的 iii),该式和(87)式等价. 因此,定理的前半部分得到证明. 若注意到(84)式和(92)完全是同一个式子, $\tilde{\boldsymbol{A}}-\tilde{\boldsymbol{B}}\tilde{\boldsymbol{B}}^T\boldsymbol{P}=\boldsymbol{A}-\boldsymbol{B}\boldsymbol{W}_{22}^{-1}(\boldsymbol{B}^T\boldsymbol{P}+\boldsymbol{W}_{12}^T)$, 根据定理 7 的 i), ii),则定理的后半部分是显然的. 证明完毕

由该结果和本节之初讲过的(84)式的解是线性矩阵不等式(86)的解的性质可以看到,若$(\boldsymbol{A}, \boldsymbol{B})$可控,则不等式条件(87)实际上也是(86)式具有实对称解的充分条件. 反之也可以证明它是必要条件.

(问题 4) 试证明上述结论(提示: 注意,若令(86)式的实对称解为$\boldsymbol{P}_0$,则可分解成$\boldsymbol{F}(\boldsymbol{P}_0; \boldsymbol{A}, \boldsymbol{B}, \boldsymbol{W})=[\boldsymbol{C}:\boldsymbol{D}]^T[\boldsymbol{C}:\boldsymbol{D}]$,利用(91)式即可)

因而,得如下定理.

[定理 10] 设$(\boldsymbol{A}, \boldsymbol{B})$可控,这时线性矩阵不等式(86)的实对称解存在的充分必要条件是(87)式成立.

(注意)

在线性矩阵不等式(86)的实对称解中, $\operatorname{rank}\boldsymbol{F}(\boldsymbol{P}; \boldsymbol{A}, \boldsymbol{B}, \boldsymbol{W})$最小,即代数方程式(84)的解必定存在,这一点一般不是自明的. 但是,若$(\boldsymbol{A}, \boldsymbol{B})$可控,则由定理 9, 10 可以得出必然存在这种解的结论(要注意,两个定理的条件一致). 而且实际上可以证明,若$(\boldsymbol{A}, \boldsymbol{B})$可稳定,线性矩阵不等式具有实对称解,则 i)必然存在最大解和最小解; ii)它们和代数方程式的解一致; iii)这些解可以作为与代数方程式相应的微分方程式对于某个起始值的定常解求得,而且该起始值本身也可以求出. 但是,由于本书篇幅的限制,故从略.

卡尔曼-亚库博维奇(Kalman-Yacubovich)引理

与下一章(A-I-17)将要讲的稳定性理论有关,有下列卡尔曼-亚库博维奇引理

"当矩阵$\boldsymbol{A}\in R^{n\times n}$(其中$\operatorname{Re}\lambda(\boldsymbol{A})<0$)和向量$\boldsymbol{b}\in R^n$(其中$(\boldsymbol{A}, \boldsymbol{b})$可控), $\boldsymbol{k}\in R^n$及正的实数γ给定时,满足联立矩阵方程式

$$\boldsymbol{P}\boldsymbol{A}+\boldsymbol{A}^T\boldsymbol{P}=-\boldsymbol{h}\boldsymbol{h}^T,\ \boldsymbol{P}\boldsymbol{b}-\boldsymbol{k}=\sqrt{\gamma}\boldsymbol{h} \tag{93}$$

的实对称矩阵$\boldsymbol{P}\in R^{n\times n}$和向量$\boldsymbol{h}\in R^n$存在的充分必要条件是,不等式条件

$$\frac{1}{2}\gamma+\operatorname{Re}\{\boldsymbol{k}^T(i\omega\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{b}\}\geqslant 0 \quad \text{a. e. } \omega\in R \tag{94}$$

成立."实际上,它是前面讲过的定理 9 的特殊情况. 下面我们先来讨论一下.

首先,若消去$\boldsymbol{h}$,将$-\boldsymbol{P}$重新写成$\boldsymbol{P}$,则(93)式和代数方程式

$$\boldsymbol{P}\boldsymbol{A}+\boldsymbol{A}^T\boldsymbol{P}-(\boldsymbol{P}\boldsymbol{b}+\boldsymbol{k})\gamma^{-1}(\boldsymbol{P}\boldsymbol{b}+\boldsymbol{k})^T=\boldsymbol{0}$$

等价. 注意到这一点,特别是在定理 9 中若设

$$\boldsymbol{B}=\boldsymbol{b},\ \boldsymbol{W}_{11}=0,\ \boldsymbol{W}_{12}=\boldsymbol{k},\ \boldsymbol{W}_{22}=\gamma,$$

则立即可以得到上述性质.

（注意）

与代数方程式(84)及线性矩阵不等式(86)的实对称解的存在有关，不等式条件(87)所起的重要作用在定理 9，10 中已证明．下面举两个例题来说明这个不等式条件的意义．

i）考虑以 p 为未知复变量的二次方程式

$$b^2p^2-2ap-c=0,$$

根据定理 9，该方程式具有实数解的充分必要条件是

$$1+\frac{b^2c}{(-i\omega-a)(i\omega-a)}=1+\frac{b^2c}{a^2+\omega^2}\geqslant 0 \quad \text{a. e. } \omega\in R$$

当然象预想那样，它和二次次方程式具有实根的充分必要条件，大家熟知的判别式条件

$$a^2+b^2c\geqslant 0$$

是等价的．由此可见，(61)式具有扩大上述判别式条件的意义．

ii）在卡尔曼-亚库博维奇引理中，若令

$$\boldsymbol{W}(s)\triangleq\frac{1}{2}\gamma+\boldsymbol{k}^T(s\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{b} \quad (s\in C),$$

因 $\boldsymbol{W}(s)$ 是 s 的实系数有理函数，而且 $\mathrm{Re}\lambda(\boldsymbol{A})<0$，则在 $\mathrm{Re}\,s>0$ 不具有极点．在这种情况下容易证明，前面的不等式条件(94)，作为 $\boldsymbol{W}(s)$ 是正实函数的充分必要条件，和大家在回路理论中所熟知的下列条件等价：

$$\mathrm{Re}\,\boldsymbol{W}(s)=\frac{1}{2}\gamma+\mathrm{Re}\{\boldsymbol{k}^T(s\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{b}\}\geqslant 0 \quad \forall\mathrm{Re}\,s>0$$

以上考察了其 2，其 3 中的黎卡提型非线性方程式，要讲的内容很多，由于篇幅的限制，下面只列举出与其有关的文献．

1）对于在其 2 中讲过的黎卡提型代数方程式

$$\boldsymbol{PA}+\boldsymbol{A}^T\boldsymbol{P}-\boldsymbol{PBB}^T\boldsymbol{P}+\boldsymbol{C}^T\boldsymbol{C}=\boldsymbol{0}, \tag{41}$$

将哈密顿矩阵定义成

$$\boldsymbol{H}\triangleq\begin{bmatrix}\boldsymbol{A} & -\boldsymbol{BB}^T\\ -\boldsymbol{C}^T\boldsymbol{C} & -\boldsymbol{A}^T\end{bmatrix}\in R^{2n\times 2n},$$

若其特征向量可以写成

$$\begin{bmatrix}\boldsymbol{x}_i\\ \cdots\cdots\\ \boldsymbol{y}_i\end{bmatrix}\begin{matrix}\}\,n\\ \\ \}\,n\end{matrix} \qquad i=1,\ 2,\ \cdots,\ n$$

则(41)式的解都可以表示成下列形式

$$\boldsymbol{P}=\boldsymbol{Y}\boldsymbol{X}^{-1}$$

式中 $\boldsymbol{X}\triangleq[\boldsymbol{x}_1,\ \boldsymbol{x}_2,\ \cdots,\ \boldsymbol{x}_n]$，$\boldsymbol{Y}\triangleq[\boldsymbol{y}_1,\ \boldsymbol{y}_2,\ \cdots,\ \boldsymbol{y}_n]$．对此，详见文献[140～143]．

2）在其 2 的定理 5 中已证明，若$(\boldsymbol{A},\ \boldsymbol{B})$可稳定，$(\boldsymbol{C},\ \boldsymbol{A})$可检出，则(41)式的唯一准正定解存在．而且，若设其为 $\boldsymbol{P}$，则

$$\mathrm{Re}\,\lambda(\boldsymbol{A}-\boldsymbol{BB}^T\boldsymbol{P})<0 \tag{43}$$

其逆，即当(41)式的唯一准正定解存在，且令其为 $\boldsymbol{P}$ 时，若(43)式成立，则$(\boldsymbol{A},\ \boldsymbol{B})$可稳定，$(\boldsymbol{C},\ \boldsymbol{A})$可检出．已经知道这个性质也成立[141, 144]．由此显而易见，在$(\boldsymbol{C},\ \boldsymbol{A})$不是可检测的情况下，(41)式的准正定解一般不是唯一的．关于这种情况下准正定解的分类及其意义，可以利用前述的哈密顿矩阵进行分析[140, 142]．

3) 在其 2 中讲过离散型黎卡提代数方程式. 这里并没有讨论将其推广到其 3 的程度,而在文献[145～146]中讨论了推广到 $\boldsymbol{W}_{22}$ 不是正定的情况.

和连续型的情况不同,在离散型的黎卡提方程式中,即使 $\boldsymbol{W}_{22}$ 不是正定的,但是也有 $(\boldsymbol{W}_{22}+\boldsymbol{B}^T\boldsymbol{P}\boldsymbol{B})$ 为正定的这样一种特殊情况. 在文献[145]中特别讨论了这种情况. 而且为此又引入新的比以前的可观测性稍强的"完备可观测性"(perfect observability)的概念. 可以证明,它对于黎卡提方程式解的存在起着重要作用. 之后在文献[146]中将该结果推广,讨论了将 $(\boldsymbol{W}_{22}+\boldsymbol{B}^T\boldsymbol{P}\boldsymbol{B})^{-1}$ 用其伪逆矩阵(参看 B-I-14)置换后的黎卡提方程式. 为此,又定义了比以前的可检测性稍强的"完备可检测性"(perfect detectability)的新的概念. 可以证明,它对于解的存在起着重要的作用.

4) 到此为止,我们讨论了方程式的性质. 对于其应用,已经几次涉及到最优控制问题[67, 116, 120]. 此外,还有估计问题[147]. 它们都以 $\boldsymbol{Q}$ 是准正定的情况为中心. 与此相反,$\boldsymbol{Q}$ 是准负定的情况,对于稳定问题[118]及谱分解[148~153]是很重要的. 特别是,线性矩阵不等式和谱分解有着密切关系. 对此可参阅文献[148～150]等等.

附录

[A] (63)式的证明

首先,令

$$\boldsymbol{D}(i\omega)\triangleq\boldsymbol{I}+\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P}_0(i\omega\boldsymbol{I}-\boldsymbol{A})^{-1}$$
$$\boldsymbol{E}(i\omega)\triangleq\boldsymbol{I}+\boldsymbol{B}^T\boldsymbol{P}_0(i\omega\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B},$$

因 $\boldsymbol{D}(i\omega)\boldsymbol{B}=\boldsymbol{B}\boldsymbol{E}(i\omega)$, 即

$$\boldsymbol{D}(i\omega)^{-1}\boldsymbol{B}=\boldsymbol{B}\boldsymbol{E}(i\omega)^{-1} \tag{A.1}$$

成立,则

$$\begin{aligned}\boldsymbol{C}_2(i\omega\boldsymbol{I}-\boldsymbol{A}_1)^{-1}\boldsymbol{B}&=\boldsymbol{C}_2[\boldsymbol{D}(i\omega)(i\omega\boldsymbol{I}-\boldsymbol{A})]^{-1}\boldsymbol{B}\\&=\boldsymbol{C}_2(i\omega\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{D}(i\omega)^{-1}\boldsymbol{B}\\&=\boldsymbol{C}_2(i\omega\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}\boldsymbol{E}(i\omega)^{-1}\end{aligned} \tag{A.2}$$

而且,根据 $\boldsymbol{P}_0$ 的定义和(62)式,则

$$\boldsymbol{E}(-i\omega)^T\boldsymbol{E}(i\omega)=\boldsymbol{I}+\boldsymbol{B}^T(-i\omega\boldsymbol{I}-\boldsymbol{A}^T)^{-1}\boldsymbol{C}_1^T\boldsymbol{C}_1(i\omega\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}\quad \text{a. e. } \omega\in R \tag{A.3}$$

成立. 将(A.2)代入(63)式最左边,利用(A.3),则

$$\begin{aligned}&\boldsymbol{I}-\boldsymbol{B}^T(-i\omega\boldsymbol{I}-\boldsymbol{A}_1^T)^{-1}\boldsymbol{C}^T\boldsymbol{C}(i\omega\boldsymbol{I}-\boldsymbol{A}_1)^{-1}\boldsymbol{B}\\&=\boldsymbol{I}-\boldsymbol{E}(-i\omega)^{-1T}\boldsymbol{B}^T(-i\omega\boldsymbol{I}-\boldsymbol{A}^T)^{-1}\boldsymbol{C}_2^T\boldsymbol{C}_2(i\omega\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}\boldsymbol{E}(i\omega)^{-1}\\&=\boldsymbol{E}(-i\omega)^{-1T}[\boldsymbol{E}(-i\omega)^T\boldsymbol{E}(i\omega)-\boldsymbol{B}^T(-i\omega\boldsymbol{I}-\boldsymbol{A}^T)^{-1}\boldsymbol{C}_2^T\boldsymbol{C}_2(i\omega\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}]\boldsymbol{E}(i\omega)^{-1}\\&=\boldsymbol{E}(-i\omega)^{-1T}[\boldsymbol{I}+\boldsymbol{B}^T(-i\omega\boldsymbol{I}-\boldsymbol{A}^T)(\boldsymbol{C}_1^T\boldsymbol{C}_1-\boldsymbol{C}_2^T\boldsymbol{C}_2)(i\omega\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}]\boldsymbol{E}(i\omega)^{-1}\\&=\boldsymbol{E}(-i\omega)^{-1T}\boldsymbol{H}(-i\omega,\ i\omega)\boldsymbol{E}(i\omega)^{-1}\quad \text{a. e. } \omega\in R\end{aligned} \tag{A.4}$$

成立,而且根据假定,因 $\boldsymbol{H}(-i\omega,\ i\omega)$ 是准正定埃尔米特矩阵,则该式的最右边也是准正定埃尔米特矩阵.

[B] (79)式的证明

根据定理 8, 设 $\boldsymbol{P}_s\triangleq\boldsymbol{P}_0-\boldsymbol{P}_+$ 时,证明

$$\|\boldsymbol{P}_s\|<\varepsilon\Rightarrow\boldsymbol{P}_0-\boldsymbol{P}_->\boldsymbol{0} \tag{B.1}$$

即可. 在这里,因矩阵 $\boldsymbol{P}_0-\boldsymbol{P}_-$ 是实对称的,则下列等值关系

$$\boldsymbol{P}_0-\boldsymbol{P}_->\mathbf{0}\Leftrightarrow\lambda\min(\boldsymbol{P}_0-\boldsymbol{P}_-)>0 \tag{B.2}$$

成立(参照 B-I-11 中第 223 页关于实对称矩阵的系 19)，而且，因

$$\begin{aligned}\lambda\min(\boldsymbol{P}_0-\boldsymbol{P}_-)&=\lambda\min[(\boldsymbol{P}_+-\boldsymbol{P}_-)+\boldsymbol{P}_\varepsilon]\\&\geqslant\lambda\min(\boldsymbol{P}_+-\boldsymbol{P}_-)+\lambda\min\boldsymbol{P}_\varepsilon\,(\text{参照 B-I-11 中第 229 页(2)'式})\\&\geqslant\lambda\min(\boldsymbol{P}_+-\boldsymbol{P}_-)-\max_i|\lambda_i(\boldsymbol{P}_\varepsilon)|\,(\text{参照 B-I-11 中第 227 页注意(i)})\\&\geqslant\lambda\min(\boldsymbol{P}_+-\boldsymbol{P}_-)-\|\boldsymbol{P}_\varepsilon\|\,(\text{参照 B-I-11 中第 226 页定理 12})\\&>\lambda\min(\boldsymbol{P}_+-\boldsymbol{P}_-)-\varepsilon\end{aligned} \tag{B.3}$$

成立，例如若选取 $\varepsilon=\lambda\min(\boldsymbol{P}_+-\boldsymbol{P}_-)\,(>0)$，由(B.1)，(B.3)，则(B.1)成立.

对于(79b)式也是同样.

[C] 首先，由定理 5 的 ii)和引理 7 的注意(i)，显然 $\boldsymbol{P}_+>\mathbf{0}$. 如引理 8 的 vi)的证明中所示，首先方程式(54)的最小解是负定的，由此及引理 7 的注意(i)中所讲过的最小解的性质，$\boldsymbol{P}_-<\mathbf{0}$ 是显然的.

[D] 若$(\boldsymbol{A}, \boldsymbol{B})$可稳定，$(\boldsymbol{C}, \boldsymbol{A})$可检测，根据定理 5，(41)式的准正定解唯一存在. 而且，若令其为 $\boldsymbol{P}$，则

$$\operatorname{Re}\lambda(\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}^T\boldsymbol{P})<0 \tag{D.1}$$

$$\boldsymbol{P}=\lim_{t\to\infty}\boldsymbol{P}(t;0,\mathbf{0}) \tag{D.2}$$

成立. 其次，设 $\boldsymbol{K}\triangleq-\boldsymbol{B}^T\boldsymbol{P}$，考虑微分方程式

$$-\dot{\tilde{\boldsymbol{P}}}=\tilde{\boldsymbol{P}}(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K})+(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{K})^T\tilde{\boldsymbol{P}}+\boldsymbol{Q}+\boldsymbol{K}^T\boldsymbol{K},\ \tilde{\boldsymbol{P}}(0)=\boldsymbol{P}_0$$

由(D.1)及引理 3，对于任意的 $\boldsymbol{P}_0=\boldsymbol{P}_0^T\geqslant\mathbf{0}$，该微分方程式具有定常解 $\lim\limits_{t\to-\infty}\tilde{\boldsymbol{P}}(t;0,\boldsymbol{P}_0)$. 而且，由(D.2)式、$\boldsymbol{K}$ 的定义及引理 5 的 ii)，该解等于(41)式的唯一准正定解 $\boldsymbol{P}$，即

$$\lim_{t\to-\infty}\tilde{\boldsymbol{P}}(t;0,\boldsymbol{P}_0)=\boldsymbol{P}\quad\forall\boldsymbol{P}_0\geqslant\mathbf{0} \tag{D.3}$$

成立. 而且，由引理 5 的(vi)，(viii)，

$$\boldsymbol{P}(t;0,\mathbf{0})\leqslant\boldsymbol{P}(t;0,\boldsymbol{P}_0)\leqslant\tilde{\boldsymbol{P}}(t;0,\boldsymbol{P}_0)\quad\forall t<0,\ \forall\boldsymbol{P}_0\geqslant\mathbf{0}$$

成立. 将该结果与(D.2)，(D.3)合起来，则可得到(81)式.

A-I-17 系统的稳定性

稳定性的概念

所谓微分方程式 $\dot{\boldsymbol{y}}=\boldsymbol{F}(\boldsymbol{y},\ t)$ 的解 $\boldsymbol{y}_c(t)$ 是稳定的，粗略地讲，是指从 $(t_0,\ \boldsymbol{y}_c(t_0))$ 的附近出发的解，在 $t\geqslant t_0$ 时充分接近 $\boldsymbol{y}_c$ 解的轨线．但是在非自治微分方程式中，根据起始时间 t_0 和起始条件 $\boldsymbol{y}_0$，解 $\boldsymbol{y}(t,\ t_0,\ \boldsymbol{y}_0)$ 的形状变化很大．与此相应地，稳定性的概念已详细定义[155~157]．那么，若进行变量变换，令 $\boldsymbol{x}\triangleq\boldsymbol{y}-\boldsymbol{y}_c$，则 $\boldsymbol{y}_c$ 的稳定性问题便归结为方程式

$$\dot{\boldsymbol{x}}=\boldsymbol{f}(\boldsymbol{x},\ t),\ \boldsymbol{f}(\boldsymbol{0},\ t)=\boldsymbol{0} \tag{1}$$

的自明解 $\boldsymbol{x}(t)\equiv 0$ 的稳定性问题．

以下假定，$\boldsymbol{f}(\boldsymbol{x},\ t)$；$R^n\times(-\infty,\ \infty)\to R^n$，(1)式对于所有的 $(t_0,\ \boldsymbol{x}_0)$ 具有唯一解，而且具有能够保证关于起始条件的解的连续性1)程度的光滑度．

［定义 1］　对于 $\forall\varepsilon>0$，$\forall t_0$，都存在 $\delta(t_0,\ \varepsilon)$，使得 $\|\boldsymbol{x}_0\|<\delta(t_0,\ \varepsilon)\Rightarrow\|\boldsymbol{x}(t,\ t_0,\ \boldsymbol{x}_0)\|<\varepsilon$，$\forall t\geqslant t_0$，这时称为 $\boldsymbol{x}(t)\equiv\boldsymbol{0}$ 在李亚普诺夫意义上是稳定的．

［定义 2］　当定义 1 中的 δ 选择成与 t_0 无关时，称为 $\boldsymbol{x}(t)\equiv\boldsymbol{0}$ 一致稳定．

［定义 3］　当(i)稳定，(ii)存在某个 $\delta_0(t_0)>0$，使得 $\|\boldsymbol{x}_0\|<\delta_0(t_0)\Rightarrow\lim\limits_{t\to\infty}\|\boldsymbol{x}(t,\ t_0,\ \boldsymbol{x}_0)\|=0$ 时，称为 $\boldsymbol{x}(t)\equiv\boldsymbol{0}$ 渐近稳定．

［定义 4］　当(i)一致稳定，(ii)有某个 $\delta_0>0$ 存在，$\|\boldsymbol{x}_0\|<\delta_0\Rightarrow$ 对于 $\forall\varepsilon>0$，有使 $\|\boldsymbol{x}(t,\ t_0,\ \boldsymbol{x}_0)\|<\varepsilon$，$\forall t,\ t_0,\ t\geqslant t_0+T$ 的 $T=T(\varepsilon)>0$ 存在时，称为 $\boldsymbol{x}(t)\equiv\boldsymbol{0}$ 一致渐近稳定．

在定义 4 中，对于 t_0，$\boldsymbol{x}_0$，要求解以同样的程度收敛于 $\boldsymbol{0}$．

因以上条件只要在 $\boldsymbol{x}(t)\equiv\boldsymbol{0}$ 的附近成立即可，所以是局部稳定的概念．大范围稳定性的定义如下．

［定义 5］　(i)稳定，(ii)在 $t\to+\infty$ 时所有的解都收敛于 $\boldsymbol{0}$，这时称 $\boldsymbol{x}(t)\equiv\boldsymbol{0}$ 大范围渐近稳定．

［定义 6］　(i)一致稳定，(ii) $\|\boldsymbol{x}_0\|<\gamma\Rightarrow\|\boldsymbol{x}(t,\ t_0,\ \boldsymbol{x}_0)\|<B(\gamma)$，$\forall\gamma>0$，(iii)对于 $\forall\alpha>0$，若 $\|\boldsymbol{x}_0\|<\alpha$，则对于 $\forall\varepsilon>0$，有使得 $\|\boldsymbol{x}(t,t_0,\boldsymbol{x}_0)\|<\varepsilon$，$\forall t,t_0,t\geqslant t_0+T$ 的 $T=T(\varepsilon,\alpha)>0$ 存在，这时称 $\boldsymbol{x}(t)\equiv\boldsymbol{0}$ 为大范围一致渐近稳定．

［定义 7］　对于 $\forall\alpha>0$，若 $\|\boldsymbol{x}_0\|<\alpha$，则 $\|\boldsymbol{x}(t,\ t_0,\ \boldsymbol{x}_0)\|\leqslant K(\alpha)\exp[-c(t-t_0)]$，$\forall t,\ t_0\ (t\geqslant t_0)$，$c>0$．这时称 $\boldsymbol{x}(t)\equiv\boldsymbol{0}$ 为大范围按指数渐近稳定．

稳定、渐近稳定、大范围渐近稳定定义之间的差别，如图 1 所示．

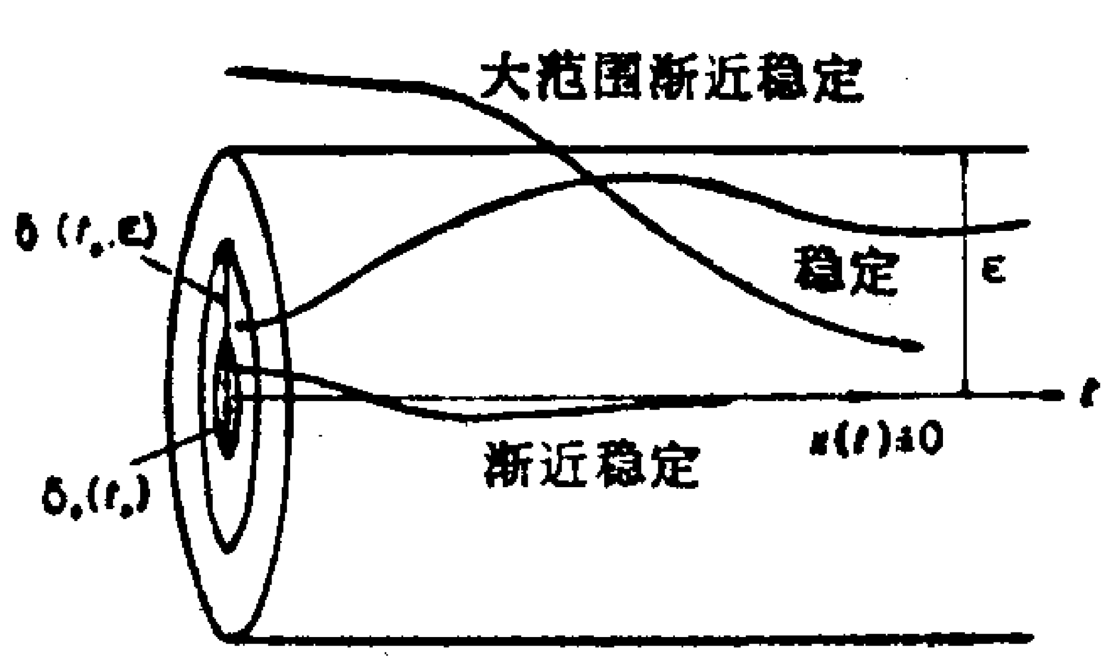


图 1　稳定的定义

1) 当 $\forall\varepsilon>0$，$\exists\delta(\varepsilon,\ t_0,\ T)>0$；$|\boldsymbol{x}_1-\boldsymbol{x}_2|<\delta(\varepsilon,\ t_0,\ T)\Rightarrow|\boldsymbol{x}(t,\ t_0,\ \boldsymbol{x}_1)-\boldsymbol{x}(t,\ t_0,\ \boldsymbol{x}_2)|<\varepsilon$，$\forall t\in[t_0,\ T]$ 时，称之为……．

以上关于稳定性的定义，起初为了明确起见给出向量的范数，但是在有限维的情况下，由于范数的等价性(B-I-5)，由定义可见，是否稳定与范数的选择无关，而仅决定于微分方程式的形式.

显然，在以上各定义之间有如下包含关系

$$
\begin{array}{ccccccc}
 & & & \nearrow & \text{定义 }2 & \searrow & \\
\text{定义 }7 \Rightarrow \text{定义 }6 \Rightarrow & \text{定义 }4 & \Rightarrow & & \text{定义 }3 & \Rightarrow & \text{定义 }1 \\
 & \searrow & \text{定义 }5 & \nearrow & & &
\end{array}
\tag{2}
$$

在线性微分方程式的情况下，得到如下等价表示.

[定理 1] 设 $\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}$ ($\boldsymbol{A}(t)$ 是由分段连续的实函数组成 n 阶方阵的状态转移矩阵为 $\boldsymbol{\Phi}(t,\tau)$.

(i) 稳定$\Leftrightarrow\|\boldsymbol{\Phi}(t,t_0)\|\leqslant N(t_0)\quad {}^\forall t\geqslant t_0,\ {}^\forall t_0$

(ii) 一致稳定$\Leftrightarrow\|\boldsymbol{\Phi}(t,t_0)\|\leqslant N,\ {}^\forall t\geqslant t_0,\ {}^\forall t_0$

(iii) 渐近稳定$\Leftrightarrow\|\boldsymbol{\Phi}(t,t_0)\|\to 0,\ t\to\infty,\ {}^\forall t_0$

(iv) 一致渐近稳定$\Leftrightarrow\|\boldsymbol{\Phi}(t,t_0)\|\leqslant N\exp[-c(t-t_0)],\ c>0,\ {}^\forall t\geqslant t_0,\ {}^\forall t_0$

(证明) 由于矩阵范数的等价性(B-I-5)，考虑导出范数就足够了. 因(i)，(ii)，(iii)的证明很容易，故从略，下面仅证明(iv). $\Leftarrow$很明显，现来证明$\Rightarrow$. 根据定义 4(i)和(ii)，有使 $\|\boldsymbol{\Phi}(t,t_0)\|\leqslant N_0,\ {}^\forall t\geqslant{}^\forall t_0,\ \|\boldsymbol{\Phi}(t,t_0)\boldsymbol{x}_0\|\leqslant 1/2,\ {}^\forall t\geqslant t_0+T,\ {}^\forall\boldsymbol{x}_0,\ \|\boldsymbol{x}_0\|=1$ 成立的 $N_0\geqslant 1,\ T>0$ 存在. 因 $\|\boldsymbol{\Phi}(t,t_0)\|=\max\limits_{\|\boldsymbol{x}_0\|=1}\|\boldsymbol{\Phi}(t,t_0)\boldsymbol{x}_0\|\leqslant 1/2,\ {}^\forall t-t_0\geqslant T$，根据转移矩阵的性质(A-I-6 中定理 5(i))，$\|\boldsymbol{\Phi}(t_0+kT,t_0)\|\leqslant N/2^k,\ k=0,1,2,\cdots$，对于任意的 $t\geqslant t_0$，若令 $kT\leqslant t-t_0<(k+1)T$，则$\|\boldsymbol{\Phi}(t,t_0)\|\leqslant\|\boldsymbol{\Phi}(t,t_0+kT)\|\|\boldsymbol{\Phi}(t_0+kT,t_0)\|\leqslant N_0^2/2^k\leqslant N\exp[-c(t-t_0)]$，其中 $N=2N_0^2,\ c=\log 2/T$.

证明完毕

(问题 1) 试证明(i)，(ii)，(iii). 提示：对于(i)，(ii)可利用 $\|\boldsymbol{\Phi}(t,t_0)\boldsymbol{x}_0\|<\varepsilon,\ \|\boldsymbol{x}_0\|=\delta$，以及 $\|\boldsymbol{\Phi}\|=\max\limits_{\|\boldsymbol{x}_0\|=\delta}\|\boldsymbol{\Phi}(t,t_0)\boldsymbol{x}_0/\delta\|$. 对于(iii)可以利用下列结果：因 $\boldsymbol{\Phi}(t,t_0)$ 关于 t 连续，而且 $\lim\limits_{t\to\infty}\|\boldsymbol{\Phi}(t,t_0)\|=0$，则 $\boldsymbol{\Phi}(t,t_0)$ 关于 t 一致有界.

(问题 2) 试利用定理 1 和(2)式来证明，在线性变系数微分方程中，定义 3$\Leftrightarrow$定义 5，定义 4$\Leftrightarrow$定义 6$\Leftrightarrow$定义 7.

[定理 2] $\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}$($\boldsymbol{A}$ 为 $n\times n$ 阶实定矩阵)

(i) 稳定$\Leftrightarrow\|\exp(\boldsymbol{A}t)\|\leqslant N,\ {}^\forall t\geqslant 0\Leftrightarrow\boldsymbol{A}$ 的最小多项式的零点全部在复平面的左半闭平面内，虚轴上零点的重数最多是 1.

(ii) 渐近稳定$\Leftrightarrow\|\exp(\boldsymbol{A}t)\|\leqslant N\exp(-ct),\ t\geqslant 0,\ c>0\Leftrightarrow\boldsymbol{A}$ 的特征值都在复平面的左半开平面内.

(证明) (i)的前半部分，由定理 1(i)和 $\boldsymbol{\Phi}(t,t_0)=\exp\boldsymbol{A}(t-t_0)$ 显然. (ii)的前半部分，由定理 1(iii)和 $\lim\limits_{t\to\infty}\|\boldsymbol{\Phi}(t,t_0)\|=0\Leftrightarrow{}^\forall\varepsilon,\ {}^\exists T(\varepsilon),\ \|\exp[\boldsymbol{A}(t-t_0)]\|<\varepsilon,\ t-t_0\geqslant T\Leftrightarrow\|\exp(\boldsymbol{A}t)\|\leqslant N\exp[-ct]$ 显然. 下面来证明(i)，(ii)的后半部分. 设 $\boldsymbol{A}$ 的最小多项式为 $\psi_m(\lambda)=(\lambda-\lambda_1)^{d_1}\cdots(\lambda-\lambda_s)^{d_s}$，其若当标准形为 $\boldsymbol{J}=\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T},\ e^{\boldsymbol{A}t}=\boldsymbol{T}e^{\boldsymbol{J}t}\boldsymbol{T}^{-1},\ e^{\boldsymbol{J}t}=\{e^{J_it}\}^{s_i}$

$=1$, $e^{J_{ij}t}=\{e^{J_{ij}t}\}^{\alpha_i}j=1$, 其中

$$\exp \mathbf{J}_{ij}t=\begin{bmatrix} e^{\lambda_i t} & te^{\lambda_i t} & \cdots & \dfrac{t^{(n_{ij}-1)}}{(n_{ij}-1)!}e^{\lambda_i t} \\ & \ddots & & \\ & & \ddots & \\ & & & e^{\lambda_i t} \end{bmatrix}=n_{ij}\times n_{ij} \tag{3}$$

式中$\{\cdot\}$表示分块对角形矩阵(B-I-13 中例 3), $\max\limits_{j} n_{ij}=d_i$(B-I-12 中定理 4). 因此, $\|e^{At}\|\leqslant N$, $t\geqslant 0\Leftrightarrow\|e^{Jt}\|\leqslant N$, $t\geqslant 0\Leftrightarrow\|e^{J_{ij}t}\|\leqslant N$, $t\geqslant 0$, $j=1, \cdots, \alpha_i$, $i=1, 2, \cdots, s\Leftrightarrow \mathrm{Re}\,\lambda_i<0(d_i\geqslant 2)$, 而且 $\mathrm{Re}\,\lambda_i=0(d_i=1)$. (ii)的后半部分证明也完全同样. 证明完毕

李亚普诺夫定理

若系统内部的能量随着时间的增加而减少, 则早晚要达到静止状态. 将这个思想发展所得到的稳定性判据, 称为李亚普诺夫第二方法(直接法). 该方法的优点是, 它对于大范围稳定性问题也有效, 而且应用它时不需要求解.

现在假定, $V(t, \boldsymbol{x})$是在 $t\in(-\infty, \infty)$, $\|\boldsymbol{x}\|<H$ 上定义的连续可微纯量函数. $V(t, \mathbf{0})=0$, $a(\|\boldsymbol{x}\|)\leqslant V(t, \boldsymbol{x})$, $^\forall\boldsymbol{x}$, $\|\boldsymbol{x}\|<H$. $a(\cdot)$是连续、单调增长, 正定, 而且 $a(0)=0$ 的函数(以下将具有该性质的函数族用 CIP 表示). 若这个$V(t, \boldsymbol{x})$ 沿(1)式解 $\boldsymbol{x}(t, t_0, \boldsymbol{x}_0)$对时间的微分 $dV(t, \boldsymbol{x}(t, t_0, \boldsymbol{x}_0))/dt$, 对于所有满足 $\|\boldsymbol{x}(t, t_0, \boldsymbol{x}_0)\|<H$ 的解是非正的, 则 $\boldsymbol{x}(t)\equiv\mathbf{0}$ 在李亚普诺夫意义上是稳定的. 其原因是, 因 $V(t, \boldsymbol{x}(t, t_0, \boldsymbol{x}_0))$ 单调非增大, 则 $a(\|\boldsymbol{x}(t, t_0, \boldsymbol{x}_0)\|)\leqslant V(t_0, \boldsymbol{x}_0)$. 因此, 对于 $^\forall\varepsilon>0$, 只要 $\boldsymbol{x}_0$ 取得足够小, $\|\boldsymbol{x}_0\|<\delta(t_0, \varepsilon)$, 则有 $V(t_0, \boldsymbol{x}_0)<a(\varepsilon)$. 对于这样的 $\boldsymbol{x}_0$, 由 $\|\boldsymbol{x}(t, t_0, \boldsymbol{x}_0)\|\leqslant a^{-1}(V(t_0, \boldsymbol{x}_0))<\varepsilon$ 便得出.

实际上, 根据定义计算 $V(t, \boldsymbol{x}(t, t_0, \boldsymbol{x}_0))$对时间的微分, 可给出如下

$$\frac{d}{dt}V(t, \boldsymbol{x}(t, t_0, \boldsymbol{x}_0))=D_{(1)}V(t, \boldsymbol{x})\big|_{\boldsymbol{x}=\boldsymbol{x}(t, t_0, \boldsymbol{x}_0)} \tag{4}$$

注意, $D_{(1)}V(t, \boldsymbol{x})=\partial V/\partial t+\boldsymbol{f}^{\mathrm{T}}\,\mathrm{grad}\,V$, 这个值, 不需要求解就可以算出这个值来。

若 $D_{(1)}V(t, \boldsymbol{x})$ 对于所有的$(t, \boldsymbol{x})$均非正, 当然 $V(t, \boldsymbol{x})$ 沿任意解对时间的微分也非正. 由以上讨论显然可得下列结果.

[定理 3]　对于某个 $H>0$, 若有在 $t\in(-\infty, \infty)$, $\|\boldsymbol{x}\|<H$ 上定义的满足下列三个条件的连续可微纯量函数 $V(t, \boldsymbol{x})$存在, 则(1)式解 $\boldsymbol{x}(t)\equiv\mathbf{0}$ 在李亚普诺夫意义上是稳定的. (i)$V(t, \mathbf{0})\equiv 0$, (ii)$a(\|\boldsymbol{x}\|)\leqslant V(t, \boldsymbol{x})$, $^\forall\boldsymbol{x}$, $\|\boldsymbol{x}\|<H$, $^\forall t$, $a(\cdot)\in$ CIP, (iii)$D_{(1)}V(t, \boldsymbol{x})\leqslant 0$, $^\forall\boldsymbol{x}$, $\|\boldsymbol{x}\|<H$, $^\forall t$.

几乎同样, 可以得到保证各种稳定性的李亚普诺夫定理. 其详细证明可参阅文献[155～157], 这里只给出结果. 以后假定, $V(t, \boldsymbol{x})$(它们称为李亚普诺夫函数)在所讨论的范围内连续可微.

[定理 4][155]　(i) $a(\|\boldsymbol{x}\|)\leqslant V(t, \boldsymbol{x})\leqslant b(\|\boldsymbol{x}\|)$, $^\forall t$, $^\forall\boldsymbol{x}$, $\|\boldsymbol{x}\|<H$, a, $b\in$ CIP; (ii) $D_{(1)}V(t, \boldsymbol{x})\leqslant 0$, $^\forall t$, $^\forall\boldsymbol{x}$, $\|\boldsymbol{x}\|<H\Rightarrow$一致稳定.

[定理 5][156]　在定理 3 的条件中, 仅将条件(iii)用 $D_{(1)}V(t, \boldsymbol{x})\leqslant -c(V(t, \boldsymbol{x}))$, $^\forall t$, $^\forall\boldsymbol{x}$, $\|\boldsymbol{x}\|<H$, $c(\cdot)\in$ CP(连续、正定, $c(0)=0$)代替$\Rightarrow$渐近稳定.

[定理 6][155] 在定理 4 的条件中，将(ii)用 $D_{(1)}V(t,\boldsymbol{x})\leqslant -c(\|\boldsymbol{x}\|)$，$^\forall t$，$^\forall\boldsymbol{x}$，$\|\boldsymbol{x}\|<H$，$c(\cdot)\in CP$ 代替⇒一致渐近稳定.

[定理 7][155] 在定理 6 的条件下，假定 $a(r)$，$b(r)\to\infty$，当 $r\to+\infty$，$H\to+\infty$⇒大范围一致渐近稳定.

[定理 8][155] (i) $\|\boldsymbol{x}\|\leqslant V(t,\ \boldsymbol{x})\leqslant\boldsymbol{K}(\alpha)\|\boldsymbol{x}\|$，$^\forall\alpha>0$，$^\forall t$，$^\forall\boldsymbol{x}$，$\|\boldsymbol{x}\|<\alpha$，(ii) $D_{(1)}V(t,\ \boldsymbol{x})\leqslant -cV(t,\ \boldsymbol{x})$，$c>0$，$^\forall t$，$^\forall\boldsymbol{x}$⇒大范围按指数渐近稳定.

(注) 在定理 7 中，$a(r)\to\infty$ 是保证解不发散的条件.

(例 1) 设 $\boldsymbol{A}(t)$ 为由分段连续一致有界的实函数组成的 $n\times n$ 阶方阵，考虑下列方程式

$$\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x} \tag{5}$$

[定理 9] 下列三个命题是等价的. (a) (5)式的自明解 $\boldsymbol{x}(t)\equiv\boldsymbol{0}$ 是一致渐近稳定的. (b)有满足以下性质的连续可微二次型李亚普诺夫函数 $V(t,\ \boldsymbol{x})=\boldsymbol{x}^T\boldsymbol{P}(t)\boldsymbol{x}$ 存在: (i) $\alpha_1\|\boldsymbol{x}\|^2\leqslant V(t,\ \boldsymbol{x})\leqslant\alpha_2\|\boldsymbol{x}\|^2$，(ii) $D_{(5)}V(t,\ \boldsymbol{x})\leqslant -c\|\boldsymbol{x}\|^2$，$c>0$. (c)对于任意的实对称、一致有界、一致正定的 $\boldsymbol{W}(t)$ ($^\exists\beta_1$，$\beta_2>0$，$\beta_1\boldsymbol{I}\leqslant\boldsymbol{W}(t)\leqslant\beta_2\boldsymbol{I}$，$^\forall t$)，下式

$$-\dot{\boldsymbol{P}}=\boldsymbol{P}\boldsymbol{A}(t)+\boldsymbol{A}^T(t)\boldsymbol{P}+\boldsymbol{W}(t) \tag{6}$$

具有和 $\boldsymbol{W}(t)$ 相同性质的解 $\boldsymbol{P}(t)$ ($^\exists\gamma_1$，$\gamma_2>0$，$\gamma_1\boldsymbol{I}\leqslant\boldsymbol{P}(t)\leqslant\gamma_2\boldsymbol{I}$，$^\forall t$).

(证明) (c)⇒(b)，令 $V(t,\ \boldsymbol{x})=\boldsymbol{x}^T\boldsymbol{P}(t)\boldsymbol{x}$ 即可. (b)⇒(a)，由定理 6 显然. 下面来证明(a)⇒(c)，先来证明

$$\boldsymbol{P}(t)=\int_t^\infty\boldsymbol{\Phi}(\lambda,\ t)^T\boldsymbol{W}(\lambda)\boldsymbol{\Phi}(\lambda,\ t)d\lambda,\ ^\forall t \tag{7}$$

存在. 由定理 1(iv)，$\|\boldsymbol{\Phi}(\lambda,\ t)\|\leqslant N\cdot\exp[-c(\lambda-t)]$，$^\forall\lambda$，$^\forall t$，$\lambda\geqslant t$. 因此，若将(7)式积分上限取为 $T(T\geqslant t)$，设这时所得结果为 $\boldsymbol{P}_T(t)$，则 $\|\boldsymbol{P}_T(t)\|\leqslant\beta_2N^2/2c$，$^\forall T\geqslant t$. 同时，若考虑将 t 固定，因 $\boldsymbol{P}_T(t)$ 对于 T 单调增大，则 $\lim\limits_{T\to\infty}\boldsymbol{P}_T(t)$ 存在(B-I-11 中定理 19)，而且 $\boldsymbol{P}(t)\leqslant\gamma_2\boldsymbol{I}$，$\gamma_2=\beta_2N^2/2c$ 成立. 再来证明 $\boldsymbol{P}(t)\geqslant\gamma_1\boldsymbol{I}$. 因 $\boldsymbol{\Phi}(\lambda,\ t)\boldsymbol{x}_0-\boldsymbol{x}_0=\int_t^\lambda\boldsymbol{A}(u)\boldsymbol{\Phi}(u,\ t)\boldsymbol{x}_0du$，根据 $\boldsymbol{A}(t)$ 的有界性和由其导出的 $\|\boldsymbol{\Phi}(u,\ t)\|\leqslant\exp[K(u-t)]$ (A-I-6 中定理 5(vi))，得 $\|\boldsymbol{\Phi}(\lambda,\ t)\boldsymbol{x}_0-\boldsymbol{x}_0\|\leqslant\omega(\lambda-t)\|\boldsymbol{x}_0\|$，$\omega(\gamma)\to 0$，$\gamma\to 0$. 因此，若将 $\sigma>0$ 选得足够小，则可得 $\|\boldsymbol{\Phi}(\lambda,\ t)\boldsymbol{x}_0\|\geqslant\|\boldsymbol{x}_0\|-\|\boldsymbol{\Phi}(\lambda,\ t)\boldsymbol{x}_0-\boldsymbol{x}_0\|\geqslant\|\boldsymbol{x}_0/2\|$，$0\leqslant\lambda-t\leqslant\sigma$.

因此

$$\begin{aligned}\boldsymbol{x}_0^T\boldsymbol{P}(t)\boldsymbol{x}_0&\geqslant\boldsymbol{x}_0^T\left(\int_t^{t+\sigma}\boldsymbol{\Phi}(\lambda,\ t)^T\boldsymbol{W}(\lambda)\boldsymbol{\Phi}(\lambda,\ t)d\lambda\right)\boldsymbol{x}_0\\&\geqslant\beta_1\int_t^{t+\sigma}\|\boldsymbol{\Phi}(\lambda,\ t)\boldsymbol{x}_0\|^2d\lambda\geqslant\beta_1\sigma\|\boldsymbol{x}_0\|^2/4\triangleq\gamma_1\|\boldsymbol{x}_0\|^2\end{aligned} \tag{8}$$

证明完毕

(注) 最后的证明部分表明，当 $\boldsymbol{A}(t)$ 有界，$\boldsymbol{0}<\beta_1\boldsymbol{I}\leqslant\boldsymbol{W}(t)=\boldsymbol{C}^T(t)\boldsymbol{C}(t)\leqslant\beta_2\boldsymbol{I}$ 时，$(\boldsymbol{A}(t),\ \boldsymbol{C}(t))$ 一致完全可观测(参照 A-I-8 引理 1).

(例 2) 在本例题中，不直接利用李亚普诺夫定理，而利用矩阵的测度 $\mu[\boldsymbol{A}]$(B-I-5)给出(5)式的稳定性判据. 设 $V(\boldsymbol{x})=\|\boldsymbol{x}\|$，则

$$D^+_{(5)}V(\boldsymbol{x})\triangleq\lim_{h\to 0_+}\frac{\|\boldsymbol{x}+h\boldsymbol{A}\boldsymbol{x}\|-\|\boldsymbol{x}\|}{h}\leqslant\lim_{h\to 0_+}\frac{\|\boldsymbol{I}+h\boldsymbol{A}\|-1}{h}\|\boldsymbol{x}\|=\mu[\boldsymbol{A}]V(\boldsymbol{x}) \tag{9}$$

左边 $D^+_{(5)}V(\boldsymbol{x})$ 可以和矩阵测度的存在完全相同地予以证明. 设 $V(\boldsymbol{x})$ 沿解的微分为

$$\frac{d}{dt}V(\boldsymbol{x}(t,\ t_0,\ \boldsymbol{x}_0))\triangleq\lim_{h\to 0_+}\frac{V(\boldsymbol{x}(t+h)-V(\boldsymbol{x}(t))}{h}$$

则

$$\frac{d}{dt}V(\boldsymbol{x}(t,\ t_0,\ \boldsymbol{x}_0))=D^+_{(5)}V(\boldsymbol{x})\big|_{\boldsymbol{x}=\boldsymbol{x}(t)}$$

几乎对所有的 t 均成立[54]. 因此, 由(9)式可得

$$V(\boldsymbol{x}(t))\leqslant V(\boldsymbol{x}_0)\exp\left[\int_{t_0}^{t}\mu[\boldsymbol{A}(s)]ds\right]$$

即

$$\|\boldsymbol{\Phi}(t,\ t_0)\|\leqslant\exp\left[\int_{t_0}^{t}\mu[\boldsymbol{A}(s)]\,ds\right].$$

由此及定理 1 可得下列结果.

[定理 10]　设 $\boldsymbol{A}(t)$ 为分段连续的实矩阵

(i) $\lim\limits_{t\to\infty}\int_{t_0}^{t}\mu[\boldsymbol{A}(s)]ds<\infty$, ${}^\forall t_0\Rightarrow$稳定

(ii) $\mu[\boldsymbol{A}(t)]\leqslant 0$, ${}^\forall t\Rightarrow$一致稳定

(iii) $\lim\limits_{t\to\infty}\int_{t_0}^{t}\mu[\boldsymbol{A}(s)]ds=-\infty$, ${}^\forall t_0\Rightarrow$渐近稳定

(iv) $\mu[\boldsymbol{A}(t)]\leqslant -a<0$, ${}^\forall t\Rightarrow$一致渐近稳定

扩大了的李亚普诺夫定理

在前面讲过的李亚普诺夫定理中, 为使解渐近地收敛到 $\mathbf{0}$, 曾付以条件 $D_{(1)}V(t,\ \boldsymbol{x})<0$, $\boldsymbol{x}\neq\mathbf{0}$ (定理 7, 8). 将该条件放宽, 拉萨尔 (Lasalle) 详细研究过 $dV(t,\ \boldsymbol{x}(t,t_0,\ \boldsymbol{x}_0))/dt\leqslant 0$ 的情况. 其结果明显表示出, 保证大范围渐近稳定的李亚普诺夫函数, 可以由更广泛的函数族中选取[158~159].

这里我们来讨论, 对于下列自由方程扩大了的李亚普诺夫定理

$$\dot{\boldsymbol{x}}=\boldsymbol{f}(\boldsymbol{x}) \tag{10}$$

其中假定, $\boldsymbol{f}(\cdot):R^n\to R^n$, 其光滑程度能够保证解的存在和唯一性, 解对于起始条件的连续性.

通过点 $\boldsymbol{x}_0$ 的(10)式解表示成 $\boldsymbol{x}(t,\ \boldsymbol{x}_0)$, 用 $\gamma(\boldsymbol{x}_0)=\{\boldsymbol{x};\boldsymbol{x}=\boldsymbol{x}(t,\ \boldsymbol{x}_0),\ -\infty<t<\infty\}$ 表示 $\boldsymbol{x}(t,\ \boldsymbol{x}_0)$ 的轨线.

所谓 $\gamma(\boldsymbol{x}_0)$ 的 ω-极限集合 $\omega(\gamma(\boldsymbol{x}_0))$, 是指当 $t\to\infty$ 时 $\boldsymbol{x}(t,\ \boldsymbol{x}_0)$ 变化到目标的集合, 即若 $\boldsymbol{q}\in\omega(\gamma)$, 则存在着某个 $t_k\to\infty$, $k\to\infty$, $\boldsymbol{x}(t_k,\ \boldsymbol{x}_0)\to\boldsymbol{q}$, $k\to\infty$ 的集合. 所谓(10)式的不变集合 M 是指, 对于 ${}^\forall\boldsymbol{p}\in M$, $\boldsymbol{x}(t,\ \boldsymbol{p})\in M$, $-\infty<t<\infty$ 时的集合.

(注)　$\omega(\gamma)$ 是不变集合. 其原因是, 对于 $\boldsymbol{p}\in\gamma$, ${}^\forall\boldsymbol{q}\in\omega(\gamma)$ 变成 $\boldsymbol{x}(t_k,\boldsymbol{p})\to\boldsymbol{q}$, $t_k\to\infty$, 对于任意的 t, $\boldsymbol{x}(t+t_k,\ \boldsymbol{p})=\boldsymbol{x}(t,\boldsymbol{x}(t_k,\boldsymbol{p}))\to\boldsymbol{x}(t,\ \boldsymbol{q})$, $t_k\to\infty$ (对于起始条件的连续性), 即 $\boldsymbol{x}(t,\ \boldsymbol{q})\in\omega(\gamma)$, ${}^\forall t$.

[定理 11]　设 $V(\boldsymbol{x})$ 为在开集 $G\subset R^n$ 上定义的可微纯量函数, 则

(i) $D_{(10)}V(\boldsymbol{x})=\boldsymbol{f}^T(\boldsymbol{x})\operatorname{grad}V(\boldsymbol{x})\leqslant 0$, ${}^\forall\boldsymbol{x}\in G$

(ii) 当 $t\geqslant 0$ 时, $\boldsymbol{x}(t,\ \boldsymbol{x}_0)$ 在 G 内且有界$\Rightarrow\boldsymbol{x}(t,\ \boldsymbol{x}_0)\to M$, $t\to\infty$.

其中 M 是包含在集合 $S=\{\boldsymbol{x};\boldsymbol{x}\in\bar{G},\ D_{(10)}V(\boldsymbol{x})=0\}$ 中的最大不变集合. $\boldsymbol{x}(t,\ \boldsymbol{x}_0)\to\boldsymbol{M}$ 表示,对于某个 $\boldsymbol{y}\in M$, 有 $\|\boldsymbol{x}(t,\ \boldsymbol{x}_0)-\boldsymbol{y}\|\to 0,\ t\to\infty$.

(证明) 因 $\omega(\gamma(\boldsymbol{x}_0))$ 是不变集合, 只要证明 $\omega\subset S$ 就足够了. 根据 (i), (ii), $\lim\limits_{t\to\infty}V(\boldsymbol{x}(t,\ \boldsymbol{x}_0))=c$ 存在. 因 $V(\cdot)$ 是连续的, 则对于 $^\forall\boldsymbol{q}\in\omega$, $V(\boldsymbol{q})=\lim\limits_{k\to\infty}V(\boldsymbol{x}(t_k,\ \boldsymbol{x}_0))=c$. 因 ω 是不变集合,则 $\boldsymbol{q}=\boldsymbol{x}(t,\ \boldsymbol{q})$, $^\forall t$, 即 $V(\boldsymbol{x}(t,\ \boldsymbol{q}))=c$, $^\forall t$, 也就是说,

$$D_{(10)}V(\boldsymbol{q})=D_{(10)}V(\boldsymbol{x})\,|_{\boldsymbol{x}=\boldsymbol{x}(t,\boldsymbol{q})}=\frac{d}{dt}V(\boldsymbol{x}(t,\ \boldsymbol{q}))\equiv 0,$$

这表示 $\boldsymbol{q}\in S$.

证明完毕

[系 1] 设 $V(\boldsymbol{x})$ 是可微的纯量函数, $G\triangleq\{\boldsymbol{x};V(\boldsymbol{x})<\rho\}$ 为有界集合. $D_{(10)}V(\boldsymbol{x})\leqslant 0$, $^\forall\boldsymbol{x}\in G\Rightarrow$对于 $^\forall\boldsymbol{x}_0\in G$, $\boldsymbol{x}(t,\ \boldsymbol{x}_0)\to M,\ t\to\infty$.

[系 2] $V(\boldsymbol{x})$ 是可微的纯量函数, (i) $V(\boldsymbol{x})\to\infty$, $\|\boldsymbol{x}\|\to\infty$, (ii) $D_{(10)}V(\boldsymbol{x})\leqslant 0$, $^\forall\boldsymbol{x}\in R^n\Rightarrow$对于 $^\forall\boldsymbol{x}_0\in R^n$, $\boldsymbol{x}(t,\ \boldsymbol{x}_0)\to M,\ t\to\infty$.

(问题 3) 试证明系 1, 2.

在以上结果中,若令 $M=\{\boldsymbol{0}\}$, 则表示 $\boldsymbol{x}(t,\ \boldsymbol{x}_0)\to\boldsymbol{0},\ t\to\infty$, 但是李亚普诺夫稳定性不清楚.

[定理 12] 设 $V(\boldsymbol{x})$ 是在 R^n 上定义的可微纯量函数, $V(\boldsymbol{0})=0$, $M=\{\boldsymbol{0}\}$.

(i) $V(\boldsymbol{x})>0$, $^\forall\boldsymbol{x}$, $\boldsymbol{x}\neq\boldsymbol{0}$

(ii) $D_{(10)}V(\boldsymbol{x})\leqslant 0$, $^\forall\boldsymbol{x}$

(iii) $V(\boldsymbol{x})\to+\infty$, $\|\boldsymbol{x}\|\to\infty\Rightarrow$大范围渐近稳定

(证明) 首先, 在系 2 中已证明, 所有的解都收敛到 $\boldsymbol{0}$. 现在来证明它是稳定的. 因 $V(\cdot)$ 连续且 $V(\boldsymbol{0})=0$, 由 (i), $\tilde{a}(\gamma)\triangleq\min\limits_{\|\boldsymbol{x}\|=\gamma}V(\boldsymbol{x})$, $\tilde{a}(\gamma)>0$, $\gamma>0$; $\tilde{a}(0)=0$. 因此, 由于 $a(\gamma)\leqslant\tilde{a}(\gamma)$, $0\leqslant\gamma<H$, $a\in$ CIP 存在, 根据定理 1, 是稳定的.

证明完毕

(注) 除了零解以外, 若 $dV(\boldsymbol{x}(t,\ \boldsymbol{x}_0))/dt$ 不恒等于 $\boldsymbol{0}$, 则 $M=\{\boldsymbol{0}\}$. 而且, 若 $D_{(10)}V(\boldsymbol{x})<0$, $^\forall\boldsymbol{x}$, $\boldsymbol{x}\neq\boldsymbol{0}$, 则 $M=\{\boldsymbol{0}\}$ (试证明之).

(例 3) 在 B-I-15 定理 1 中已证明, $\boldsymbol{A}$ 渐近稳定$\Leftrightarrow\boldsymbol{PA}+\boldsymbol{A}^T\boldsymbol{P}+\boldsymbol{L}^T\boldsymbol{L}=\boldsymbol{0}$ 对于任意 $(\boldsymbol{A},\ \boldsymbol{L})$ 可观测的 $\boldsymbol{L}$, $\boldsymbol{P}=\boldsymbol{P}^T>\boldsymbol{0}$. 现在利用定理 12 来证明$\Leftarrow$. 设 $V(\boldsymbol{x})=\boldsymbol{x}^T\boldsymbol{P}\boldsymbol{x}$, 求沿 $\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}$ 的解 $\boldsymbol{x}(t,\ \boldsymbol{x}_0)$ 对时间的微分, 得 $dV(\boldsymbol{x}(t,\ \boldsymbol{x}_0))/dt=-\boldsymbol{x}(t,\ \boldsymbol{x}_0)^T\boldsymbol{L}^T\boldsymbol{L}\boldsymbol{x}(t,\ \boldsymbol{x}_0)$. 由此可见, V 完全满足定理 12 的条件. 下面来证明 $M=\{\boldsymbol{0}\}$.

$dV(\boldsymbol{x}(t,\ \boldsymbol{x}_0))/dt\equiv 0\Leftrightarrow\boldsymbol{L}\exp[\boldsymbol{A}t]\boldsymbol{x}_0\equiv\boldsymbol{0}\Leftrightarrow\boldsymbol{x}_0=\boldsymbol{0}$ (由可观测性), 由此及定理 12 的 (注), 得 $M=\{\boldsymbol{0}\}$.

(例 4) 考虑下列自治方程[160~163]

$$\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}\varphi(\sigma)\quad \sigma=-\boldsymbol{c}^T\boldsymbol{x} \tag{11}$$

式中 $\boldsymbol{A}\in R^{n\times n}$, $\boldsymbol{b}\in R^{n\times 1}$, $\boldsymbol{c}\in R^{n\times 1}$, $\mathrm{Re}\lambda[\boldsymbol{A}]<0$, 假定 $(\boldsymbol{A},\boldsymbol{b},\boldsymbol{c})$ 可控可观测, 而且 $\varphi(\cdot)\in A_k$, 设

$$A_k=\{\varphi;\varphi\text{ 连续},\ \varphi(0)=0,\ 0\leqslant\varphi/\sigma<k,\ ^\forall\sigma\neq 0,$$
$$\int_0^\sigma\varphi(u)du\to\infty,\ |\sigma|\to\infty\} \tag{12}$$

当(11)式的零解对于所有的 $\varphi(\cdot)\in A_k$ 都大范围渐近稳定时，称其为绝对稳定.

(11)式是表示包含非线性 $\varphi(\cdot)$ 的反馈系统方程式.

[定理 13]　设 $(\boldsymbol{A},\ \boldsymbol{b},\ \boldsymbol{c})$ 可控可观测，$\mathrm{Re}\lambda[\boldsymbol{A}]<0$. 这时，(11)，(12)式的零解绝对稳定的充分条件是下列矩阵方程式

$$\boldsymbol{PA}+\boldsymbol{A}^T\boldsymbol{P}=-\boldsymbol{q}\boldsymbol{q}^T \tag{13}$$

$$\boldsymbol{Pb}-(\boldsymbol{c}+\theta\boldsymbol{A}^T\boldsymbol{c})=\sqrt{\gamma}\,\boldsymbol{q} \tag{14}$$

式中

$$\gamma=2(k^{-1}+\theta\boldsymbol{c}^T\boldsymbol{b})\geqslant 0 \tag{15}$$

对于满足(15)式的某个实数 $\theta\geqslant 0$，具有实数解 $\boldsymbol{P}=\boldsymbol{P}^T\in R^{n\times n}$，$\boldsymbol{q}\in R^{n\times 1}$.

为了证明该定理，先来证明下列引理.

[引理 1]　设定理 13 的条件由 $\boldsymbol{P}$，$\boldsymbol{q}$，$\theta\geqslant 0$ 满足，则 $\boldsymbol{P}=\boldsymbol{P}^T\geqslant\boldsymbol{0}$，特别是

(ⅰ) 当 $\theta=0$ 时，

$$\boldsymbol{P}>\boldsymbol{0} \tag{16}$$

(ⅱ) 当 $\theta>0$ 时，

$$\boldsymbol{x}_0^T\boldsymbol{P}\boldsymbol{x}_0>0,\ \forall\boldsymbol{x}_0\in\{\boldsymbol{x};\ \boldsymbol{x}\neq\boldsymbol{0},\ \boldsymbol{c}^T\boldsymbol{x}=0\} \tag{17}$$

成立.

(证明)　由(13)式给出

$$\boldsymbol{P}=\int_0^{\infty}e^{\boldsymbol{A}^Tt}\boldsymbol{q}\boldsymbol{q}^Te^{\boldsymbol{A}t}dt \tag{18}$$

显然，$\boldsymbol{P}=\boldsymbol{P}^T\geqslant\boldsymbol{0}$.

现在来证明(ⅰ)，(ⅱ). 首先注意，若假定 $\boldsymbol{x}_0^T\boldsymbol{P}\boldsymbol{x}_0=0$，则由(18)式得 $\boldsymbol{q}^T\exp[\boldsymbol{A}t]\boldsymbol{x}_0\equiv 0$. 因此，由(13)式，$d\boldsymbol{x}_0^Te^{\boldsymbol{A}^Tt}\boldsymbol{P}e^{\boldsymbol{A}t}\boldsymbol{x}_0/dt=-\boldsymbol{x}_0^Te^{\boldsymbol{A}^Tt}\boldsymbol{q}\boldsymbol{q}^Te^{\boldsymbol{A}t}\boldsymbol{x}_0\equiv 0$，即 $\boldsymbol{x}_0^Te^{\boldsymbol{A}^Tt}\boldsymbol{P}e^{\boldsymbol{A}t}\boldsymbol{x}_0=\boldsymbol{x}_0^T\boldsymbol{P}\boldsymbol{x}_0=0$. 由这些结果及(14)式可得，$(\boldsymbol{c}+\theta\boldsymbol{A}^T\boldsymbol{c})^Te^{\boldsymbol{A}t}\boldsymbol{x}_0=(\boldsymbol{b}^T\boldsymbol{P}-\sqrt{\gamma}\,\boldsymbol{q}^T)e^{\boldsymbol{A}t}\boldsymbol{x}_0\equiv 0$. 因此，设 $\theta=0$，若考虑到可观测性，则 $\boldsymbol{x}_0=\boldsymbol{0}$，即(ⅰ)得到证明. (ⅱ)，若设 $\boldsymbol{c}^T\exp[\boldsymbol{A}t]\boldsymbol{x}_0=y(t)$，则 $\theta\dot{y}+y=0$，即

$$y(t)=y(0)\exp[-t/\theta]=\boldsymbol{c}^T\boldsymbol{x}_0\exp[-t/\theta]\equiv 0$$

根据上面结果及可观测性，则 $\boldsymbol{x}_0=\boldsymbol{0}$，(17)式得到证明.

证明完毕

(定理 13 的证明)　当 $\varphi(\cdot)\equiv 0$ 时，大范围渐近稳定是显然的. 因此，以下 $\forall\varphi\in A_k$ 考虑为 $\varphi\neq 0$. 设 $\boldsymbol{P}$，$\boldsymbol{q}$，θ 为按定理 13 的条件求出的，考虑下列李亚普诺夫函数

$$V(\boldsymbol{x})=\boldsymbol{x}^T\boldsymbol{P}\boldsymbol{x}+2\theta\int_0^{\sigma}\varphi(u)du,\ \sigma=-\boldsymbol{c}^T\boldsymbol{x} \tag{19}$$

显然，$V(\cdot)$ 连续可微，$V(\boldsymbol{0})=0$. 下面来检证定理 12 的条件(ⅰ)～(ⅲ).

(ⅰ) $V(\boldsymbol{x})>0$，$\forall\boldsymbol{x}$，$\boldsymbol{x}\neq\boldsymbol{0}$：当 $\theta=0$ 时，由引理 1(ⅰ)显然. 假定 $\theta>0$，由

$$\boldsymbol{P}\geqslant 0,\ \theta\int_0^{\sigma}\varphi(u)du\geqslant 0$$

显然 $V(\boldsymbol{x})\geqslant 0$(因为根据(12)式条件，$\varphi(\cdot)$ 可用 Ⅰ，Ⅲ 象限内的图形表示). 因此，现在对于 $\varphi\neq 0$ 和 $\boldsymbol{x}_0\neq\boldsymbol{0}$，若设 $V(\boldsymbol{x}_0)=0$，则由(19)式得 $\boldsymbol{x}_0^T\boldsymbol{P}\boldsymbol{x}_0=0$，$\boldsymbol{c}^T\boldsymbol{x}_0=0$. 这与引理 1 的(ⅱ)相矛盾.

(ⅱ) $D_{(11)}V(\boldsymbol{x})\leqslant 0$，$\forall\boldsymbol{x}$：利用 $D_{(11)}V(\boldsymbol{x})=\dfrac{d}{dt}V(\boldsymbol{x}(t,\ \boldsymbol{x}_0))\big|_{\boldsymbol{x}(t,\boldsymbol{x}_0)=\boldsymbol{x}}$ 计算，得

$$D_{(11)}V(\boldsymbol{x})=\boldsymbol{x}^T(\boldsymbol{PA}+\boldsymbol{A}^T\boldsymbol{P})\boldsymbol{x}+2\varphi\boldsymbol{b}^T\boldsymbol{Px}-2\theta(\mathbf{c}^T\boldsymbol{Ax}+\mathbf{c}^T\boldsymbol{b}\varphi)\boldsymbol{\varphi}$$
$$=-(\boldsymbol{q}^T\boldsymbol{x}-\sqrt{\gamma}\varphi)^2-2\left(\sigma\varphi-\frac{1}{k}\right)\varphi^2\leqslant 0 \tag{20}$$

(iii) $V(\boldsymbol{x})\to\infty$, $\|\boldsymbol{x}\|\to\infty$: 当 $\theta=0$ 时, 由 $\boldsymbol{P}>\mathbf{0}$ 显然. 设 $\theta>0$. 将 $\forall\boldsymbol{x}\neq\mathbf{0}$ 分解成 $\boldsymbol{x}=\alpha\mathbf{c}+\boldsymbol{x}_1$, $\mathbf{c}^T\boldsymbol{x}_1=0$, $\alpha=\mathbf{c}^T\boldsymbol{x}/\mathbf{c}^T\mathbf{c}$. 当 $|\alpha|\to+\infty$ 时, 由条件(12)得

$$V(\boldsymbol{x})\geqslant 2\theta\int_0^{-\alpha\mathbf{c}^T\mathbf{c}}\varphi(u)du\to+\infty.$$

当 $|\alpha|<\infty$ 时, 因 $\|\boldsymbol{x}_1\|\to\infty$, 则 $V(\boldsymbol{x})\geqslant\boldsymbol{x}_1^T\boldsymbol{P}\boldsymbol{x}_1+2\alpha\boldsymbol{x}_1^T\boldsymbol{P}\mathbf{c}\geqslant\mu\|\boldsymbol{x}_1\|^2-2|\alpha|\|\boldsymbol{x}_1\|\|\boldsymbol{P}\mathbf{c}\|\to+\infty$(由(17)式, $\mu>0$ 存在).

(iv) 最后, 来证明 $M=\{\mathbf{0}\}$. 由(20)式得,

$$\frac{d}{dt}V(\boldsymbol{x}(t,\ \boldsymbol{x}_0))\equiv 0\Rightarrow\boldsymbol{q}^T\boldsymbol{x}(t,\ \boldsymbol{x}_0)-\sqrt{\gamma}\varphi(\sigma(t))\equiv 0,$$
$$\varphi(\sigma(t))(k\sigma(t)-\varphi(\sigma(t)))\equiv 0.$$

因此, 若考虑到条件(12), 则 $\varphi(\sigma(t))\equiv 0$, $\boldsymbol{q}^T\boldsymbol{x}(t,\ \boldsymbol{x}_0)\equiv 0$, 这表明 $\boldsymbol{x}(t,\ \boldsymbol{x}_0)=\exp[\boldsymbol{A}t]\boldsymbol{x}_0$, $\boldsymbol{q}^T\exp[\boldsymbol{A}t]\boldsymbol{x}_0\equiv 0$. 根据(18)式, 显然 $\boldsymbol{x}_0^T\boldsymbol{P}\boldsymbol{x}_0=0$. 因

$$V(\boldsymbol{x}(t,\ \boldsymbol{x}_0))=\boldsymbol{x}_0^T\boldsymbol{P}\boldsymbol{x}_0+2\theta\int_0^{-\mathbf{c}^T\boldsymbol{x}_0}\varphi(u)du,\ \lim_{t\to\infty}\boldsymbol{x}(t,\ \boldsymbol{x}_0)=\lim_{t\to\infty}\exp[\boldsymbol{A}t]\boldsymbol{x}_0=0,$$

则必须

$$\boldsymbol{x}_0^T\boldsymbol{P}\boldsymbol{x}_0+2\theta\int_0^{-\mathbf{c}^T\boldsymbol{x}_0}\varphi(u)du=0.$$

因 $\varphi\neq 0$, 则 $\boldsymbol{x}_0^T\boldsymbol{P}\boldsymbol{x}_0=0$, $\mathbf{c}^T\boldsymbol{x}_0=0$. 根据引理 1(ii), 这表明 $\boldsymbol{x}_0=\mathbf{0}$, 即 $\boldsymbol{x}(t,\ \boldsymbol{x}_0)\equiv\mathbf{0}$.

证明完毕

再者, 根据卡尔曼-亚库博维奇引理[160, 161](参照上一章), 当 $\gamma>0$ 时, (13), (14)式具有实解 $\boldsymbol{P}$, $\boldsymbol{q}$ 的充分必要条件是, 对于所有的实数 ω, $(\gamma/2)+\mathrm{Re}\{(\mathbf{c}+\theta\boldsymbol{A}^T\mathbf{c})^T(i\omega\boldsymbol{I}-\boldsymbol{A}^{-1})\boldsymbol{b}\}\geqslant 0$ 成立. 该条件和

$$\frac{1}{k}+\mathrm{Re}[(1+i\omega\theta)\mathbf{c}^T(i\omega\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{b}]\geqslant 0,\ \forall\omega\in R^1$$

等价(试证明), 它是作为波波夫(popov)条件的绝对稳定性条件[163].

(注) 在前一章讲过的卡尔曼-亚库博维奇引理中, $\gamma>0$ 的条件是必要的, 但是在卡尔曼的论文中[160]证明, 即使令该条件为 $\gamma\geqslant 0$ 也成立. 因此, 定理 13 的条件和波波夫条件等价.

(注) 迈耶(Meyer)证明[162], 在例 4 所讨论的绝对稳定性问题中, 即使条件

$$\int_0^\sigma\varphi(u)du\to\infty,\ |\sigma|\to\infty$$

不成立, 但是若波波夫条件成立, 则仍是绝对稳定的.

线性系统有界输入-有界输出的稳定性

在李亚普诺夫意义上的稳定性, 是着眼于系统内部的情况, 即系统状态的行为. 这里所谓有界输入-有界输出的稳定性, 是从系统外部观察系统的行为, 即指输入输出关系的性质. 它们是两个不同的稳定性概念, 没有直接关系, 在自动控制中都很重要.

现在假定某个线性系统的输入为 $\boldsymbol{u}(t)\in R^r$, 输出为 $\boldsymbol{y}(t)\in R^m$, 脉冲响应为 $\boldsymbol{H}(t,\ \boldsymbol{\tau})$

$\in R^{m\times r}$，则在 $t=t_0$ 处于零状态的系统的输出，即零状态响应，一般可表示成[1)]

$$\boldsymbol{y}(t)=\int_{t_0}^{t}\boldsymbol{H}(t,\ \tau)\boldsymbol{u}(\tau)d\tau,\ t>t_0 \tag{21}$$

在这里，为了弄清楚"有界"的意义，将向量范数固定，使用最方便的 $\|\cdot\|_1$(B-I-5)，特别是，以后不再加脚注了，例如 $\|\boldsymbol{u}(t)\|\triangleq|u_1(t)|+|u_2(t)|+\cdots+|u_r(t)|$. 对于矩阵，使用由向量范数导出的范数($\boldsymbol{A}=[a_{ij}]$, $\|\boldsymbol{A}\|\triangleq\max\limits_{j}\sum\limits_{i}|a_{ij}|$). 在 $t\geqslant t_0$ 定义的输入是有界的是指，对于 $\boldsymbol{u}$ 确定某个常数 $u_M<\infty$, $\|\boldsymbol{u}(t)\|<u_M$, $^\forall t>t_0$ 成立.

再者，注意到(21)式规定的线性系统的零状态响应，有界输入-有界输出的稳定性定义如下.

[定义 8]　对于任意 t_0 和任意有界输入，当 $\|\boldsymbol{y}(t)\|\leqslant c(t_0,\ \boldsymbol{u})$, $^\forall t\geqslant t_0$ 对于某个常数 $c(t_0,\ \boldsymbol{u})$ 成立时[2)]，称为系统是有界输入-有界输出稳定(简写成 BIBO 稳定).

[定义 9]　当系统是 BIBO 稳定，常数 $c(t_0,\ \boldsymbol{u})$ 不依存于 t_0，即 $c=c(\boldsymbol{u})$ 时，称为系统一致 BIBO 稳定.

[定义 10]　对于任意的 t_0 和任意的有界输入，当 $\|\boldsymbol{y}(t)\|\leqslant k(t_0)u_M$, $^\forall t\geqslant t_0$ 对于某个常数 $k(t_0)$ 成立时，称为系统同程度 BIBO 稳定.

[定义 11]　当系统同程度 BIBO 稳定，常数 $k(t_0)$ 不依存于 t_0，即 $k(t_0)=k$ 时，称为系统一致同程度 BIBO 稳定.

根据以上定义，显然有下列包含关系

一致同程度 BIBO 稳定 ⇒ 同程度 BIBO 稳定 ⇒BIBO 稳定

⇘ 一致 BIBO 稳定 ⇗

(一致)　BIBO 稳定和(一致)同程度 BIBO 稳定的差别是，前者只不过是保证输入输出之间的"增益" $\|\boldsymbol{y}(t)\|/u_M$ 是有限的，而后者是具有 $k(t_0)$ (k) 的上限值，只有这个是严格的条件. 但是对于线性系统，如下面定理 14 所示，实际上它们是等价的概念，无论用哪个都一样.

[定理 14][164]　对于零状态响应按(21)式规定的线性系统，下列等价关系成立.

(i) BIBO 稳定⇔同程度 BIBO 稳定

$$\Leftrightarrow\int_{t_0}^{t}\|\boldsymbol{H}(t,\ \tau)\|d\tau\leqslant c(t_0),\ ^\forall t>t_0 \tag{22}$$

(ii) 一致 BIBO 稳定⇔一致同程度 BIBO 稳定

$$\Leftrightarrow\int_{-\infty}^{t}\|\boldsymbol{H}(t,\ \tau)\|d\tau\leqslant c,\ ^\forall t \tag{23}$$

(证明)　(i)首先注意，作为 BIBO 稳定的必要条件

$$\int_{t_0}^{t}\|\boldsymbol{H}(t,\ \tau)\|d\tau\leqslant c(t_0,\ t)<\infty\ ,\ ^\forall t_0,\ ^\forall t\geqslant t_0 \tag{24}$$

成立. 这可证明如下. 对于各 t_0, $t\geqslant t_0$，若令

$$\boldsymbol{u}(\tau)=\tilde{\boldsymbol{u}}(\tau;t,\ t_0)\triangleq\left\{\begin{array}{ll}(0,\ \cdots,\ 0,\ \overset{\text{第 } j \text{ 个成分}}{\overset{\downarrow}{\operatorname{sgn} h_{ij}(t,\ \tau)}},\ 0,\ \cdots,\ 0)^T, & t_0\leqslant\tau<t\\ \boldsymbol{0} & \tau>t\end{array}\right. \tag{25}$$

1) 假定 $\boldsymbol{H}(t,\ \tau)$ 至少为可测函数.

2) $c(t_0,\ \boldsymbol{u})$ 意味着，c 一般对应于 t_0 和输入确定.

显然，$\sup_{\tau>t_0}\|\tilde{\boldsymbol{u}}(\tau; t, t_0)\|=1$，若为 BIBO 稳定，则有与 t_0，$\boldsymbol{u}$ 对应的某个 $c_{ij}=c_{ij}(t_0, t)$ 存在，便得

$$c_{ij}(t_0, t)\geqslant\|\boldsymbol{y}(t)\|\geqslant|y_{ij}(t)|=\int_{t_0}^{t}|h_{ij}(t, \tau)|d\tau \tag{26}$$

成立．由(26)式，设 $c(t_0, t)=\sum_{i,j=1}c_{ij}(t_0, t)$，则(24)式成立．即，将 t_0，t 固定时，(24)式左边的积分是有界的．因此，若再利用所谓的共鸣定理[165]，则可以证明对于 t 一致有界，即

$$\int_{t_0}^{t}\|\boldsymbol{H}(t, \tau)\|d\tau\leqslant c(t_0) \tag{27}$$

（参照附录）．因此证明

$$\|\boldsymbol{y}(t)\|\leqslant\int_{t_0}^{t}\|\boldsymbol{H}(t, \tau)\|\|\boldsymbol{u}(\tau)\|d\tau\leqslant c(t_0)u_M \tag{28}$$

和

BIBO 稳定⇒(27)式⇒同程度 BIBO 稳定

成立．若为同程度 BIBO 稳定，根据定义，由于它也是 BIBO 稳定，则(22)式成立．

(ii) 和(24)式的推导完全一样，可以证明

$$\text{一致 BIBO 稳定}\Rightarrow\int_{-\infty}^{t}\|\boldsymbol{H}(t, \tau)\|d\tau\leqslant c(t)<\infty, \ \forall t \tag{29}$$

而且由共鸣定理可以证明，(27)式对于 $c(t_0)=c$ 成立．由以上结果显然．

（注意）

(i) 因定理 14 对于由脉冲响应 $\boldsymbol{H}(t, \tau)$ 规定零状态响应的线性(变系数)系统成立，故不受系统是有限维的限制．

(ii) 所谓 $\boldsymbol{u}(\cdot)$，$\boldsymbol{y}(\cdot)$ 是有界的，与 R^r，R^m 的范数选择无关(范数的等价性)．因此，上述输入输出稳定性是系统固有的性质．但是若有界的意义定义为能量有界，

$$\int_{t_0}^{\infty}\|\cdot\|^2dt<\infty,$$

则与上述的 BIBO 稳定完全不同(在无限维的情况下，不存在范数的等价性)．

(iii) 特别是，对于零状态响应由

$$\boldsymbol{y}(t)=\int_{t_0}^{t}\boldsymbol{H}(t-\tau)\boldsymbol{u}(\tau)d\tau, \ t\geqslant t_0 \tag{30}$$

表示的常系数系统，定义 8～11 均等价，作为定理 14 的特殊情况

[系 2]

$$\text{BIBO 稳定}\Leftrightarrow\int_{0}^{\infty}\|\boldsymbol{H}(\tau)\|d\tau\leqslant c \tag{31}$$

成立．

$\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}+\boldsymbol{B}(t)\boldsymbol{u}$，$\boldsymbol{y}=\boldsymbol{C}(t)\boldsymbol{x}$ 的 BIBS 稳定和渐近稳定的关系

首先，作为研究由外部观察的系统稳定性和内部状态稳定性的关系的第一步，考虑下列 BIBS 稳定．

[定义 12]

在具有状态方程式

$$\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}+\boldsymbol{B}(t)\boldsymbol{u}, \ \boldsymbol{x}(t_0)=\boldsymbol{x}_0$$
$$\boldsymbol{A}(t)\in R^{n\times n}, \ \boldsymbol{B}(t)\in R^{n\times r} \tag{32}$$

的系统中，对于任意的 t_0，任意的有界输入，任意的 $\boldsymbol{x}_0\in R^n$，当 $\|\boldsymbol{x}(t)\|\leqslant c(t_0, \boldsymbol{x}_0, \boldsymbol{u})<\infty$, $^\forall t\geqslant t_0$ 成立时，系统称为有界输入-有界状态稳定(BIBS 稳定).

[定义 13]　当系统是 BIBS 稳定，$c(t_0, \boldsymbol{x}_0, \boldsymbol{u})=c(\boldsymbol{x}_0, \boldsymbol{u})$ 时，称为是一致 BIBS 稳定.

[定义 14]　对于任意的 t_0，任意的有界输入，关于由 $\boldsymbol{x}(t_0)=\boldsymbol{x}_0=\boldsymbol{0}$ 出发的状态 $\boldsymbol{x}(t)$，当 $\|\boldsymbol{x}(t)\|\leqslant k(t_0)u_M$, $^\forall t\geqslant t_0$ 成立时，系统称为同程度 BIBS 稳定.

[定义 15]　当系统是同程度 BIBS 稳定，$k(t_0)=k$ 时，称为是一致 BIBS 稳定.

一般，设 $\boldsymbol{A}(\cdot)$ 的转移矩阵为 $\boldsymbol{\Phi}(\cdot, \cdot)$ 时

$$\boldsymbol{x}(t)=\boldsymbol{\Phi}(t, t_0)\boldsymbol{x}_0+\int_{t_0}^{t}\boldsymbol{\Phi}(t, \tau)\boldsymbol{B}(\tau)\boldsymbol{u}(\tau)d\tau. \tag{33}$$

对(33)式右边第 2 项应用定理 14，显然

$$\text{同程度 BIBS 稳定}\Leftrightarrow\int_{t_0}^{t}\|\boldsymbol{\Phi}(t, \tau)\boldsymbol{B}(\tau)\|d\tau\leqslant c(t_0)\quad {}^\forall t\geqslant t_0 \tag{34}$$

$$\text{一致同程度 BIBS 稳定}\Leftrightarrow\int_{-\infty}^{t}\|\boldsymbol{\Phi}(t, \tau)\boldsymbol{B}(\tau)\|d\tau\leqslant c,\ {}^\forall t \tag{35}$$

成立．此外，在 $x(t)$ 的表示式((33)式)中，根据定理 1(a)((b))，因为表示起始状态影响的右边第 1 项有界(一致有界)的充分必要条件是，$\|\boldsymbol{\Phi}(t, t_0)\|\leqslant N(t_0)\,(N)$，则将其归纳得如下定理.

[定理 15]　系统 $\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}+\boldsymbol{B}(t)\boldsymbol{u}$ 是 BIBS 稳定(一致 BIBS 稳定)的充分必要条件是，(i)系统是同程度 BIBS 稳定(一致同程度 BIBS 稳定)，(ii)稳定(一致稳定).

此外，为了推导出与输入输出响应有关的稳定性和内部状态的稳定性之间的关系，必须对系统的可控性，可观测性作一些假定．因此，以下对于所讨论的系统 $\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}+\boldsymbol{B}(t)\boldsymbol{u}$, $\boldsymbol{y}(t)=\boldsymbol{C}(t)\boldsymbol{x}$，假定为一致完全可控和一致完全可观测的，设 $\boldsymbol{A}(\cdot)$, $\boldsymbol{B}(\cdot)$, $\boldsymbol{C}(\cdot)$ 分段连续且有界，即假定 $\|\boldsymbol{A}(t)\|\leqslant K$, $\|\boldsymbol{B}(t)\|\leqslant K$, $\|\boldsymbol{C}(t)\|\leqslant K$, $^\forall t$ 成立．那么在这种情况下，一致完全可控，一致完全可观测的充分必要条件是

$$\boldsymbol{W}_c(t-\sigma, t)\triangleq\int_{t-\sigma}^{t}\boldsymbol{\Phi}(t, \tau)\boldsymbol{B}(\tau)\boldsymbol{B}^T(\tau)\boldsymbol{\Phi}^T(t, \tau)d\tau\underset{^\forall t}{\geqslant}\alpha(\sigma)\boldsymbol{I}>\boldsymbol{0}, \tag{36}$$

$$\boldsymbol{W}_0(t, t+\sigma)\triangleq\int_{t}^{t+\sigma}\boldsymbol{\Phi}^T(t, \tau)\boldsymbol{C}^T(\tau)\boldsymbol{C}(\tau)\boldsymbol{\Phi}(t, \tau)d\tau\underset{^\forall t}{\geqslant}\beta(\sigma)\boldsymbol{I}>0, \tag{37}$$

(A-I-8 中定义 6, A-I-13″ 中(13b)′ 式).

[定理 16]　在 $\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}+\boldsymbol{B}(t)\boldsymbol{u}$, $\boldsymbol{y}=\boldsymbol{C}(t)\boldsymbol{x}$ 中，设 $\|\boldsymbol{C}(t)\|\leqslant K$，而且一致可观测性((37)式)成立．这时

(i) 同程度 BIBS 稳定⇔同程度 BIBO 稳定

(ii) 一致同程度 BIBS 稳定⇔一致同程度 BIBO 稳定

(证明)　(i)⇒: 因脉冲响应为 $\boldsymbol{H}(t, \tau)=\boldsymbol{C}(t)\boldsymbol{\Phi}(t, \tau)\boldsymbol{B}(\tau)$, $\boldsymbol{C}(\cdot)$有界，由

$$\int_{t_0}^{t}\|\boldsymbol{H}(t, \tau)\|d\tau\leqslant K\int_{t_0}^{t}\|\boldsymbol{\Phi}(t, \tau)\boldsymbol{B}(\tau)\|d\tau\leqslant Kc(t_0) \tag{38}$$

成立，显然.

⇐: 对于任意的 t_0 和任意的 $\|\boldsymbol{u}(t)\|\leqslant u_M$, $^\forall t\geqslant t_0$ 的输入，设

$$\tilde{\boldsymbol{u}}_t(\tau)\triangleq\begin{cases}\boldsymbol{u}(\tau), & t_0\leqslant\tau\leqslant t\\ \boldsymbol{0}, & \tau>t.\end{cases}$$

因对应于 $\tilde{\boldsymbol{u}}_t$ 的状态在 $\tau \geqslant t$ 时可表示成 $\boldsymbol{x}(\tau)=\boldsymbol{\Phi}(\tau, t)\boldsymbol{x}(t)$，则

$$\int_t^{t+\sigma} \boldsymbol{x}^T(\tau)\boldsymbol{C}^T(\tau)\boldsymbol{C}(\tau)\boldsymbol{x}(\tau)d\tau=\boldsymbol{x}^T(t)\boldsymbol{W}_0(t, t+\sigma)\boldsymbol{x}(t)\geqslant\beta(\sigma)\boldsymbol{x}^T(t)\boldsymbol{x}(t). \tag{39}$$

此外，根据同程度 BIBO 稳定的定义，有某个 $c(t_0)$ 存在，对于 $\tilde{\boldsymbol{u}}_t$，

$$\|\boldsymbol{y}(\tau)\|=\|\boldsymbol{C}(\tau)\boldsymbol{x}(\tau)\|\leqslant c(t_0)u_M \tag{40}$$

成立．由(39)，(40)式得

$$\|\boldsymbol{x}(t)\|\leqslant(\sigma/\beta(\sigma))^{\frac{1}{2}}c(t_0)u_M. \tag{41}$$

因(41)式对于任意的 $t\geqslant t_0$ 均成立，故为同程度 BIBS 稳定．

(ii) 只要设 $c(t_0)$ 为常数 c，可以和(i)完全同样地证明．

最后，作为本节的重要结果，下面介绍一下阐明内部稳定和外部稳定关系的定理．

[定理 17][166] 在系统 $\dot{\boldsymbol{x}}=\boldsymbol{A}(t)\boldsymbol{x}+\boldsymbol{B}(t)\boldsymbol{u}$, $\boldsymbol{y}=\boldsymbol{C}(t)\boldsymbol{x}$ 中，设 $\|\boldsymbol{A}(t)\|\leqslant K$, $\|\boldsymbol{B}(t)\|\leqslant K$, $\|\boldsymbol{C}(t)\|\leqslant K$, $\forall t$，而且系统是一致完全可控，一致完全可观测的．这时

(i) (a)BIBO 稳定⇔(b)同程度 BIBS 稳定

$$\Leftrightarrow(c)\int_{t_0}^t\|\boldsymbol{\Phi}(t, \tau)\|d\tau\leqslant c(t_0),\ \forall t \Rightarrow \text{(d) 渐近稳定}$$

(ii) (a)一致 BIBO 稳定⇔(b)一致同程度 BIBS 稳定

$$\Leftrightarrow(c)\int_{t_0}^t\|\boldsymbol{\Phi}(t, \tau)\|d\tau\leqslant c,\ \forall t \Leftrightarrow \text{(d) 一致渐近稳定}$$

(证明) (i)：因(a)⇔(b)在定理 16 中已证明，则下面先证明(b)⇐(c)．因

$$\int_{t_0}^t\|\boldsymbol{\Phi}(t, \tau)\boldsymbol{B}(\tau)\|d\tau\leqslant K\int_{t_0}^t\|\boldsymbol{\Phi}(t, \tau)\|d\tau\leqslant Kc(t_0),\ \forall t$$

由(34)式可以证明(b)⇐(c)．其次，再来证明(b)⇒(c)．因 $\|\boldsymbol{A}(t)\|\leqslant K,\ \forall t \Rightarrow \|\boldsymbol{\Phi}(t, s)\|\leqslant\exp(K|t-s|)$, $\forall t, s$(参照 A-I-11 后面的附录)，则

$$\|\boldsymbol{B}^T(\tau)\boldsymbol{\Phi}^T(s, \tau)\|\leqslant nrK\exp(K\sigma)=\gamma_1(\sigma),\ \forall s,\ \forall \tau,\ s-\sigma\leqslant\tau\leqslant s \tag{42}$$

而且

$$\|\boldsymbol{W}_c^{-1}(s-\sigma, s)\|\leqslant\sqrt{n}\,\alpha(\sigma)^{-1}\triangleq\gamma_2(\sigma),\ \forall s \tag{43}$$

因此

$$\begin{aligned}
&\int_{t_0}^t\|\boldsymbol{\Phi}(t, s)\|ds\\
&=\int_{t_0}^t\|\boldsymbol{\Phi}(t, s)\boldsymbol{W}_c(s-\sigma, s)\boldsymbol{W}_c^{-1}(s-\sigma, s)\|ds\\
&\leqslant\gamma_2(\sigma)\int_{s=t_0}^t\left\|\int_{\tau=s-\sigma}^s\boldsymbol{\Phi}(t, \tau)\boldsymbol{B}(\tau)\boldsymbol{B}^T(\tau)\boldsymbol{\Phi}^T(s, \tau)\right\|d\tau\, ds\\
&\leqslant\gamma_1(\sigma)\gamma_2(\sigma)\int_{s=t_0}^t\int_{\tau=s-\sigma}^s\|\boldsymbol{\Phi}(t, \tau)\boldsymbol{B}(\tau)\|d\tau\, ds\\
&\leqslant\gamma_1(\sigma)\gamma_2(\sigma)\int_{x=0}^{\sigma}\int_{\tau=t_0+x-\sigma}^{t+x-\sigma}\|\boldsymbol{\Phi}(t, \tau)\boldsymbol{B}(\tau)\|d\tau\, dx\\
&\leqslant\gamma_1(\sigma)\gamma_2(\sigma)\sigma\int_{t_0-\sigma}^t\|\boldsymbol{\Phi}(t, \tau)\boldsymbol{B}(\tau)\|d\tau\\
&\leqslant\gamma_1(\sigma)\gamma_2(\sigma)\sigma c(t_0-\sigma)\triangleq\bar{c}(t_0)
\end{aligned} \tag{44}$$

其中最后的不等式是根据(34)式．

最后，来证明(c)⇒(d)．首先

$$\|\boldsymbol{\Phi}(t, t_0)-\boldsymbol{I}\|=\left\|\int_{t_0}^{t}\frac{d}{d\tau}\boldsymbol{\Phi}(t, \tau)d\tau\right\|$$

$$=\left\|\int_{t_0}^{t}\boldsymbol{\Phi}(t, \tau)\boldsymbol{A}(\tau)d\tau\right\|\leqslant K\int_{t_0}^{t}\|\boldsymbol{\Phi}(t, \tau)\|d\tau\leqslant Kc(t_0)$$

所以 $\|\boldsymbol{\Phi}(t, t_0)\|\leqslant 1+Kc(t_0)$, $^\forall t\geqslant t_0$ 成立. 因此

$$\begin{aligned}\|\boldsymbol{\Phi}(t, t_0)\|&=\frac{1}{t-t_0}\int_{t_0}^{t}\|\boldsymbol{\Phi}(t, t_0)\|d\tau\\&\leqslant\frac{1}{t-t_0}\int_{t_0}^{t}\|\boldsymbol{\Phi}(t, \tau)\|\|\boldsymbol{\Phi}(\tau, t_0)\|d\tau\\&\leqslant\frac{1+Kc(t_0)}{t-t_0}\int_{t_0}^{t}\|\boldsymbol{\Phi}(t, \tau)\|d\tau\\&\leqslant c(t_0)\frac{1+Kc(t_0)}{t-t_0}\to 0,\ t\to\infty\end{aligned}\tag{45}$$

由此结果及定理 1(iii) 证明是渐近稳定的.

(ii): (b)⇔(c): 只要在(i)的证明中令 $c(t_0)\to c$ 即可. (c)⇐(d) 可证明如下. 若为一致渐近稳定, 由定理 1(iv) 得

$$\int_{t_0}^{t}\|\boldsymbol{\Phi}(t, \tau)\|d\tau\leqslant N\int_{t_0}^{t}\exp(-\alpha(t-\tau))d\tau\leqslant\frac{N}{\alpha}$$

反之, 当(c)成立时, 由(45)式,

$$\|\boldsymbol{\Phi}(t, t_0)\|\leqslant\frac{c(1+Kc)}{t-t_0}\leqslant\varepsilon,\ ^\forall t\geqslant t_0+\varepsilon c(1+Kc)\tag{46}$$

成立, 根据定义 4, 则为一致渐近稳定. 即证明了(c)⇒(d).

对于常系数系统 $\dot{\boldsymbol{x}}=\boldsymbol{A}x+\boldsymbol{B}\boldsymbol{u}$, $\boldsymbol{y}=\boldsymbol{C}\boldsymbol{x}$, 一致稳定和稳定是等价的. 因此, 由(35)式得

$$\text{同程度 BIBS 稳定}\Leftrightarrow\int_{0}^{\infty}\|e^{\boldsymbol{A}\tau}\boldsymbol{B}\|d\tau\leqslant c.\tag{47}$$

由文献[28]中知道, (47)式右边的积分有界的充分必要条件是: $\boldsymbol{A}$ 仅具有 $\mathrm{Re}\lambda(\boldsymbol{A})\leqslant 0$ 的特征值, 特别是, $\mathrm{Re}\lambda(\boldsymbol{A})=0$ 的特征值是最小多项式的单根, 而且对应的非衰减系统模型不结合输入. 因此可以理解, 若 $(\boldsymbol{A}, \boldsymbol{B})$ 可控, 则不允许 $\mathrm{Re}\lambda(\boldsymbol{A})=0$, 具有全部 $\mathrm{Re}(\lambda)<0$ 的特征值是同程度 BIBS 稳定的充分必要条件.

对于常系数系统, 定理 17 变成如下形式.

[系 3] 在系统 $\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{u}$, $\boldsymbol{y}=\boldsymbol{C}\boldsymbol{x}$ 中, 设 $(\boldsymbol{A}, \boldsymbol{B})$ 可控且 $(\boldsymbol{A}, \boldsymbol{C})$ 可观测. 这时, BIBO 稳定⇔BIBS 稳定⇔同程度 BIBS 稳定⇔渐近稳定⇔$\boldsymbol{A}$ 的特征值均在复平面的左半开平面内⇔传递函数矩阵 $\boldsymbol{H}(s)=\boldsymbol{C}[s\boldsymbol{I}-\boldsymbol{A}]^{-1}\boldsymbol{B}+\boldsymbol{D}$ 的极点均在左半开平面内.

(问题 1) 系统 $\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{B}\boldsymbol{u}$, $\boldsymbol{y}=\boldsymbol{C}\boldsymbol{x}+\boldsymbol{D}\boldsymbol{u}$ 是 BIBO 稳定的充分必要条件是, 传递函数矩阵 $\boldsymbol{H}(s)=\boldsymbol{C}[s\boldsymbol{I}-\boldsymbol{A}]^{-1}\boldsymbol{B}+\boldsymbol{D}$ 在左半开平面内具有极点, 试证明之.

(问题 2) 试证明系 3.

(问题 3) 在离散时间系统 $\boldsymbol{x}(k+1)=\boldsymbol{A}\boldsymbol{x}(k)$ 中, 试证明下列关系

(i) 稳定⇔$\|\boldsymbol{A}^k\|\leqslant c$, $^\forall k$⇔$|\lambda(\boldsymbol{A})|\leqslant 1$, 而且 $|\lambda(\boldsymbol{A})|=1$ 是最小多项式的单根.

(ii) 渐近稳定⇔$\boldsymbol{A}^k\to\boldsymbol{0}$, $k\to\infty$⇔$|\lambda(\boldsymbol{A})|=1$⇔对于任意的 $\boldsymbol{Q}=\boldsymbol{Q}^T>\boldsymbol{0}$, 矩阵方程式

$A^TPA-P=Q$ 具有解 $P=P^T>0$.

附录

[共鸣定理][165] 设 X：完备的赋范空间，Y：赋范空间，T_n: $X\to Y$，$n=1,2,\cdots$ 为线性有界加法算子．对于 $^\forall x\in X$，若 $\{\|T_nx\|\}$ 均为有界数列，则 $\{\|T_n\|\}$ 也"共鸣"，为有界数列．

再者，设 $X=\{u;\sup_{\tau\geq t_0}\|u(\tau)\|_1<\infty\}$，$Y=\{y;\|y\|_1<\infty\}$ 若 X 中的范数取为

$$\sup_{\tau\geq t_0}\|u(\tau)\|_1,$$

Y 中的范数取为 $\|y\|_1$，则 X 变成巴拿赫空间，Y 为赋范空间．

在各个 $t\geq t_0$，若令

$$T_tu\triangleq\int_{t_0}^{t}H(t,\tau)u(\tau)d\tau$$

根据(24)式，则 T_t: $X\to Y$ 为线性算子，而且 $\{\|T_tu\|\}$ 为有界序列．因此，由共鸣定理得，$\|T_t\|$ 对于 t 一致有界，$\int_{t_0}^{t}\|H(t,\tau)\|d\tau\leq rm\|T_t\|\leq c(t_0)$.

参 考 文 献

[1] 高桥安人: システムと制御, 岩波書店, (昭 43).
[2] 電氣学会编: 自動制御理論, コロナ社, (昭 45).
[3] Y. Takahashi et al: Control, Addison Wesley, 1970.
[4] K. Ogata: Modern Control Engineering Prentice Hall (1970).
[5] S. C. Gupta & L. Hasdorff: Fundamentals of Automatic Control, John Wiley, (1970).
[6] 弥永昌吉等 ほか: 代数学, 岩波讲座現代応用數学, (昭 32).
[7] 弥永昌吉等 ほか: 现代數学概説 I, 岩波書店, (昭 36).
[8] G. Birkhoff & S. MacLane: A Survey of Modern Algebra, MacMillan, (1965).
[9] 斎藤正彦: 線型代數入门, 東京大学出版会, (昭 41).
[10] 佐武一郎: 行列と行列式, 裳華房, (昭 33).
[11] S. Karlin: Mathematical Methods and Theory in Games, Programming and Economics, Vol. I, II, Addison Wesley, Reading Mass, (1959).
[12] G. Hadley: Linear Algebra, Addison-Wesley, Reading, Mass, (1961).
[13] G. Hadley: Linear Programming, Addison-Wesley, Reading, Mass, (1962).
[14] M. D. Canon, C. D. Cullum & E. Polak: Theory of Optimal Control and Mathematical Programming, McGraw-Hill, New York, (1970).
[15] A. V. Fiacco & G. P. McCormik: Nonlinear Programming; Sequential Unconstrained Minimization Techniques, John Wiley, New York, (1968).
[16] O. L. Mangasarian: Nonlinear Programming, McGraw Hill, New York, (1969).
[17] A. S. Lasdon: Optimization Theory for Large Systems, MacMillan, London, (1970).
[18] 関根泰次: 數理計画法 I, II, 岩波讲座, 基礎工学, 岩波書店, (昭 43).
[19] H. G. Campbell: "Linear Algebra with Applications", Appleton Century Crofts, (1971).
[20] 古屋茂: 行列と行列式, 培風舘, (1957).
[21] M. Athans: Information and Control, II, pp. 592～606, (1967).
[22] M. Marcus & H. Ming: A Survey of Matrix Theory and Matrix Inequalities; Allyn and Bacon, Inc. (1964).
[23] C. T. Chen: Introduction to Linear System Theory; Holt, Rinehart and Winston. Inc. (1970).
[24] D. K. Faddeev & V. N. Faddeeva: Computational Methods of Linear Algebra; W. H. Freeman and Co. (1963).
[25] F. R. Gatmacher: The Theory of Matrix; Vols. 1 & 2, Chelsea, (1959).
[26] L. G. Kelly: Handbook of Numerical Methods and Applications; Addison-Wesley, (1967).
[27] E. A. Guillemin: The Mathematics of Circuit Analysis; John Wiley, (1964).
[27a] 戸川隼人: マトリクスの數值計算, オーム社, (昭 46).
[28] C. A. Desoer: Notes for Second Course on Linear Systems; Van Nostrand, (1970).
[29] 伊沢計介: 自動制御入門: オーム社, (昭 29).
[30] 渡部和: 計算機による電子回路設計概論; 電子通信学会誌, pp. 578～595, (昭 44-5).
[31] L. A. Zadeh & C. A. Desoer: Linear System Theory; McGraw-Hill, (1963).
[32] C. A. Desoer & E. S. Kuh: Basic Circuit Theory; McGraw-Hill, (1969).
[33] S. Seshu & M. B. Reed: Linear Graphs and Electrical Networks; Addison-Wesley, (1961).
[34] 伊理正夫: 線形回路網の位相幾何学的な取极い線形集中定數系論 I 第 3 章; 岩波讲座基礎工学 6, 岩波書店, (昭 44)
[35] H. R. Martens & D. R. Allen: Introduction to Systems Theory; Charles E. Merrill Publishing Co., Ohio, (1969).
[36] R. A. Rohrer: Circuit Theory; An Introduction to the State Variable Approach; McGraw-Hill, (1970).
[37] E. S. Kuh & R. A. Rohrer: The State-Variable Approach to Network Analysis; Proc. IEEE, Vol. 53, No. 7, pp. 672～682, (1965).
[38] C. Pottle: State-Space Techniques for General Active Network Analysis; System Analysis by Digital Computers, F. F. Kuo, J. F. Kaiser, John Wiley, New York, (1966).

[39] D. A. Calahan: Computer-Aided Network Design; McGraw-Hill, New York, (1968).
[40] C. Pottle: A "Textbook" Computerized State-Space Network Analysis Algorithm; IEEE Trans. Circuit Theory, Vol. CT-16, No. 4, pp. 566～568, (1969).
[41] 大越孝敬: 大学课程基礎電子回路; オーム社, 東京. (昭 42).
[42] A. Dervisoğlu: Comments on the Existence of the A State Matrix; IEEE Trans. Circuit Theory, Vol. CT-16, No. 2, p. 242, (1969).
[43] A. Dervisoğlu: State Equations and Initial Values in Active RLC Networks; IEEE Trans. Circuit Theory, Vol. CT-18, No. 5, pp. 544～547, (1971).
[44] A. Fettweis: On the Algebraic Derivation of the State Equations, IEEE Trans. Circuit Theory, Vol. CT-16, No. 2, pp. 171～175, (1969).
[45] P. H. O'N. Roe: Networks and Systems; Addison Wesley, Reading, (1966).
[46] H. E. Koenig, Y. Tokad, H. K. Kesavan & H. G. Hedges: Analysis of Discrete Physical Systems; Mc-Graw-Hill, New York, (1967).
[47] H. R. Martens & D. R. Allen: Introduction to Systems Theory; Charles E. Merrill Publishing Co., Columbus, (1969).
[48] 高桥安人: ダイナミックシステム論; 科学技術社, 金沢, (昭 45).
[49] 村上吉繁・住本彰・西村正太郎: 回路論的手法による電氣-力学系のシミュレーション; 電子通信学会回路とシステム理論研究会资料, CT71-4, (昭 46-5).
[49a] Truxal: Automatic Control System Synthesis; Chap. 2, McGraw Hill, (1955).
[50] 日本自動制御協会: プロセス制御ハンドブック; 第 14 章, 朝倉書店, (昭 46).
[51] 松原正一: プロセスシステム工学; 第 5 章, 朝倉書店, (昭 45).
[53][1] H. Haneda: Analysis of Computational Techniques for Circuit Theory; Electronics Research Laboratory Memorandom, No. ERL-M323, University of California(Berkeley), Feb. 16, (1972).
[54] W. A. Coppel: Stability and Asymptotic Behavior of Differential Equations; D. C. Heath Co., Boston, (1965).
[55] G. Dahlquist: Stability and Error Bounds in the Numerical integration of Ordinary Differential Equation; Trans. of the Royal Inst. of Tech., Stockholm, Sweden, No. 130, (1959).
[56] W. J. Vetter: Some Extensions of Matrix Methods for Large System Studies; Proc. of 14th Midwest Symp. On Circuit Theory, (1971).
[57] E. S. Kuh & I. Hajj: Nonlinear Circuit Theory; Resistive Networks, Proc. IEEE, Vol. 59, No. 3, pp. 340～355, (1971).
[58] ポントリャーギン: 常微分方程式(新版), pp. 290～300, 共立出版, (昭 43).
[59] H. D'Angelo: Linear Time-Varying Systems; Allyn and Bacon, Boston, (1970).
[60] P. Lancaster: Theory of Matries; Academic Press, (1969).
[61] K. Ogata: State Space Analysis of Control Systems; Prentice Hall, (1967).
[62] E. F. Beckenbach & R. Bellman: Inequalities; Springer-Verlag, Berlin, (1971).
[63] R. Bellman: Introduction to Matrix Analysis; McGraw-Hill, New York, (1960).
[64] K. L. Hitz et al: A Note on Bounds on Solutions of the Riccati Equation; IEEE Trans. Automatic Control, p. 178, Feb, (1972).
[65] 兒玉・须田・白川: 非線形制御系の安定判別における線形化法-II; 制御工学, 14 巻, 12 号, pp. 751～762, (昭 45-12).
[66] C. A. Desoer: A Counter-Example Useful in Stability Theory; Notes on System Theory, Vol. II, Electronics Research Laboratory, Univ. of California, Berkeley, Feb., (1962).
[67] R. E. Kalman: Contribution to the Theory of optimal control; Boletin de la Sociedad. Matematica Mexicana, Ser II, 5, (1960).
[68] M. Athans: The Design of Suboptimal Linear Time-Varying Systems; IEEE Trans. Automatic Control, AC-13, 150, (1968).
[69] R. Bellman: Some Inequalities for the Square Root of a Positive Definite Matrix; Linear Algebra and its Applications, I, pp. 321～324, (1968).

1) 原文参考文献的编号有些混乱，中间有缺号现象(译者).

[70] F. T. Man: Some Inequalities for Positive Detinite Symmetric Matrices; SIAM J. Applied Mathematics, Vol. 19, No. 4, pp. 679～681, Dec. (1970).

[71] H. H. Rosenbrock: State-Space and Multivariable Theory; Nelson, (1970).

[72] R. .ツルミール(瀬川,高市訳): マトリクスの理論と応用; ブレイン図書出版, (1972).

[73] M. L. Liou: A Novel Method of Evaluating Transient Response; Proc. IEEE, Vol. 54, pp. 20～23, Jan. (1966).

[74] S. Genapathy & A. S. Rao: Transient Response Evaluation from the State Transition Matrix; Proc. IEEE (Letters), Vol. 57, pp. 347～349, March, (1969).

[75] E. J. Mastascusa: A Method for Calculating e^{At} Based on the Cayley-Hamilon Theorem; Proc. IEEE (Letters), Vol. 57, pp. 1328～1329, July, (1969).

[76] 平野・貴田: 線形システムの過渡応答の計算時間と誤差の評価; 電子通信学会, 回路とシステム理論研究会資料, CT72-50, (昭 47-11).

[77] M. Silverberg: New Method of Solving State Variable Equations Permitting Large Step Sizes; Proc. IEEE (Letters), Vol. 56, pp. 1352～1353, Aug. (1968).

[78] W. Everling: On the Evaluation of e^{At} by Power Series; Proc. IEEE (Letters), Vol. 55, p. 413, March, (1967).

[79] T. A. Bichart: Matrix Exponential; Approximation by Truncated Power Series; Proc. IEEE (Letters), Vol, 56, pp. 872～873, May, (1968).

[80] Director & R. A. Rohrer: Introduction to System Theory; McGraw Hill, New York, (1972).

[81] P. M. Derusso, R. J. Roy & C. M. Close: State Variables for Engineers; Wiley, New York, (1965).

[82] L. M. Silverman: Realization of Linear Dynamical Systems; IEEE Trans. Automatic Control, Vol. AC-16, No. 6, pp. 554～567, Dec. (1971).

[83] 前田: 有向グラフの制御理論への応用, 計測自動制御学会論文集, 8 巻, 5 号, pp. 619～626, (昭 47-10).

[84] W. M. Wonham & C. D. Johnson: A Note on the Transformation to Canonical (phase-variable) Form; Trans. IEEE, Vol. AC-9, pp. 312～313, July, (1964).

[85] I. H. Mufti: On the Reduction of a System to Canonical (phase-variable) Form; Trans. IEEE, Vol. AC-10, pp. 206～207, April, (1965).

[86] M. R. Chidambra: The Transformation to (phase-variable) Canonical Form; Trans. IEEE, Vol. AC-10, pp. 492～495, Oct. (1965).

[87] W. G. Tuel, Jr.: On the Transformation to (phase-variable) Canonical Form; Trans. IEEE, Vol. AC-11, p. 607, July, (1966).

[88] B. D. O. Anderson & D. G. Luenberger; Design of Multivariable Feedback Systems; Proc. IEEE (London), Vol. 114, pp. 395～399, March, (1967).

[89] S. J. Asseo: Phase Variable Canonical Transformation of Multicontroller Systems; Trans. IEEE, Vol. AC-13, pp. 129～131, (1968).

[90] D. G. Luenberger: Canonical forms for linear multivariable systems; Trans. IEEE, Vol. AC-12, pp. 290～293, (1967).

[91] C. D. Johnson: A Unified Canonical Form for controllable Linear Dynamical Systems; Proc. (1969). JACC, pp. 189～199, (1969).

[92] 横山: 線形多入力多出力系の Canonical Form; 計測自動制御学会論文集, Vol. 7, No. 5, pp. 42～48, (昭 46).

[93] P. Brunovsky: A Classification of Linear Controllable Systems; Kybernetic Cislo, 3, pp. 173～188, (1970).

[94] R. E. Kalman: Kronecker Invariants and Feedback; Ordinary Differential Equations, pp. 459～471 (1972) (L. Weiss, Editor)

[95] W. M. Wonham & A. S. Morse: Feedback Invariants of Linear Multivariable Systems; Automatica, 8, pp. 93～100, (1972).

[96] V. M. Popov: Invariant Description of Linear Time-Invariant Controllable Systems; SIAM J. Control, 10, pp. 252～264, (1972).

[97] D. G. Luenberger: Observing the State of a Linear System; IEEE Trans. Mil, Electronics, Vol. MIL-8, pp. 74～80, April, (1964).

[98] T. E. Fortmann & D. Williamson: Disign of Low-Order Observers for Linear Feedback Control Laws; IEEE Trans. Automatic Control, Vol. AC-17, pp. 301~308, June, (1972).
[99] D. G. Luenberger: An Introduction to Observers; IEEE Trans. Automatic Control, Vol. AC-16, pp. 596~602, Dec. (1971).
[100] B. Gopinath: On the Control of Linear Multiple Inout-Output Systems; Bell System Tech. J. March, (1971).
[101] 前田:線形汎関数観測器の構成;システムと制御,16巻,7号,pp. 580~584,(昭 47-7).
[101a] 増淵;オブザーバと補償器;计测自動制御学会讲習会「最適化手法」テキスト,(昭 48-5).
[102] W. A. Wolovich: The Application of State Feedback Invariants to Exact Model Matching. Presented at the 5th Annual Princeton Conf. Information Science and Systems; Princeton, N. J., pp. 387~392, Mar. (1971).
[103] S. H. Wang & C. A. Desoer: The Exact Model Matching of Linear Multivariable Systems; IEEE Trans. Automat. Contr. (Short Papers), Vol. AC-17, pp. 347~349, June, (1972).
[104] B. C. Moore & L. M. Silverman: Model Matching by State Feedback and Dynamic Compensation; IEEE Trans. On Automat. Contr., Vol. AC-17, No. 4, pp. 491~497, Aug., (1972).
[105] B. S. Morgan, Jr.: The Synthesis of Linear Multi-Variable Systems by State Feedback, in Proc. Joint Automat. Contr. Conf., pp. 468~472, (1964).
[106] P. L. Falb & W. A. Wolovich: Decoupling in the Design and Synthesis of Multivariable Control Systems; IEEE Trans. Automat. Contr., Vol. AC-12, pp. 651~659, Dec, (1967).
[107] E. G. Gilbert: The Decoupling of Multivariable Systems by State Feedback; SIAM J. Contr., Vol. 7, pp. 50~63, Feb., (1969).
[108] W. M. Wonham & A. S. Morse: Decoupling and Pole-Assignment in Linear Multivariable Systems, A Geometric Approach; SIAM J. Contr., Vol. 8, pp. 1~18, Feb. (1970).
[109] A. S. Morse & W. M. Wonham: Decoupling and Pole Assignment by Dyanmic Compensation; SIAM J. Contr., Vol. 8, pp. 317~337, Aug. (1970).
[110] A. S. Morse & W. M. Wonham: Status of Noninteracting Control; IEEE Trans. On Automat. Contr., Vol. AC-16, No. 6, pp. 568~581, Dec. (1971).
[111] L. M. Silverman: Properties and Application of Inverse Systems; IEEE Trans, Automat. Contr. (Short Papers), Vol. AC-13, pp. 436~437, Aug. (1968).
[112] M. L. Sain & J. L. Massey: Invertibitity of Linear Time Invariant Dynamical Systems; IEEE Trans. Automat. Contr., Vol. AC-14, pp. 141~149, April, (1969).
[113] P. Dorato: On the Invese of Linear Dynamical Systems; IEEE Trans. Syst. Sci. Cybernetics, Vol. SSC-5, pp. 43~48, Jan. (1969).
[114] L. M. Silverman: Inversion of Multi-variable Linear Systems; IEEE Trans. Automat. Contr., Vol. AC-14, pp. 270~276, June, (1969).
[115] L. M. Silverman & H. J. Payne: Input-Output Structure of Linear Systems wish Application to the Decoupling Problem; SIAM J. Contr., Vol. 9, pp. 199~233, (1971).
[116] M. Athans & P. L. Falb: Optimal Control, An Introduction to the Theory and Its Applications; McGraw-Hill, New York, (1966).
[117] R. E. Kalman, P. L. Falb & M. Arbib: Topics in Mathematical System Theory; McGraw-Hill, New York, (1969).
[118] R. W. Brockett: Finite Dimensional Linear Systems; Wiley, New York, (1970).
[119] B. D. O. Anderson & J. B. Moore: Lineer Optimal Control; Prentice-Hall, Englewood Cliffs N. J. (1971).
[120] H. Kwakernaak & R. Sivan: Linear Optimal Control Systems; by John Wiley & Sons, Inc. (1972).
[121] R. E. Kalman: When is a Linear Control System Optimal?; Trans. ASME Ser. D. J. Basic Engrg., 86, pp. 51~60, (1964).
[122] 須賀:最適制御における逆問題;計測と制御,6巻,8号,(昭 42).
[123] E. Kreindler & J. K. Hedrick: On Equivalence of Quadratic Loss Functions; Internat. J. Control, 11, pp. 213~222, (1970).
[124] P. J. Moylan & B. D. O. Anderson: Non-Linear Regulator Theory and an Inverse Optimal Control

Problem; in Proc. Joint Automat. Contr., Conf. pp. 462～467, (1971).
[125] R. Yokoyama & E. Kinnen: The Inverse Problem of the Optimal Regulator; IEEE Trans. On Automat. Contr., Vol. AC-17, No. 4, pp. 497～504, Aug. (1972).
[126] A. Jameson & E. Kreindler: Inverse Problem of Linear Optimal Control; SIAM J. Control, Vol. 11, No. 1, pp. 1～19, Feb. (1973).
[127] 池田・前田・児玉: 線形時変システムの状態フィードバックによる設計; 電子通信学会論文誌, Vol. 55-D, No. 11, pp. 707～714, (1972).
[128] C. R. Rao & S. K. Mitra: Generalized Inverse of Matrices and Its Application; Wiley, NewYork, (1971).
[129] T. L. Boullion & P. L. Odell: Generalized Inverse Matrices; John Wiley & Sons, Inc. (1971).
[130] A. Albert: Regression and the Moore-Penrose Pseudoinverse; Academic Press, New York and London, (1972).
[131] 児玉・池田・藤田: 線形離散ダイナミカルシステムの表現式について; 第3回ダイナミカル・システムシンポジウム, 日本自動制御協会7月13,14日, (昭48).
[132] 池田・富田: 線形離時間システムに対する線形関数観測器について; 電子通信学会論文誌D, 技術談話室(採録決定).
[133] 井上・室井: 線形離散時間型制御系のオブザーバ; 第3回ダイナミカルシステムシンポジウム.
[134] 市川: 離散時間形のオブザーバ; 計測自動制御学会学術讲演会, 8月25,26,27日, (昭46).
[135] 舟橋・中村: 線形離散時間系のオブザーバ設計; 第4回统計学的制御理論シンポジウム, 11月27,28,29日, (昭47).
[136] 長田・西村・瓜倉・池田: 線形離散システムの関数オブザーバ; 計測自動制御学会学術讲演会, 8月20・21・22日, (昭48).
[137] W. M. Wonham: On a Matrix Riccati Equation of Stochastic Control; SIAM J. Contr., Vol. 14, pp. 681～698, (1960).
[138] J. C. Willems: Least squares Stationary Optimal Control and the Algebraic Riccati Equation; IEEE Trans. Automatic Contr., AC-16, No. 6, pp. 621～634, Dec. (1971).
[140] K. Martensson: On the Riccati Equation; Inform. Sci., Vol. 3, pp. 17～49, Jan. (1971).
[141] V. Kučera: A Contribution to Matrix Quadratic Equations; IEEE Trans. Automatic Contr., (Short Papers), Vol. AC-17, pp. 344～347, June, (1972).
[142] V. Kučera: On Nonnegative Definite Solutions to Matrix Quadratic Equations; Automatica, Vol. 8, No. 4, pp. 413～423, July, (1972).
[143] Stephen Barnett: Matrices in Control Theory; Van Nostrand Reinhold Company, London, (1971).
[144] H. J. Payne: An Alternate Proof Related to the Algebraic Riccati Equation; IEEE Trans. Automatic Contr., (Technical Note and Correspondence), Vol. AC-16, p. 822, Dec, (1972).
[145] D. Rappaport & L. M. Silverman: Structure and Stability of Discrete-Time Optimal Systems; IEEE Trans. Automatic Contr., Vol. AC-16, pp. 227～233, June, (1971).
[146] H. J. Payne & L. M. Silverman: On the Discrete Time Algebraic Riccati Equation; IEEE Trans. Automatic Contr., Vol. AC-18, No. 3, pp. 226～234, June, (1973).
[147] R. E. Kalman: A New Approach to Linear Filtering and Perdiction Problems; Trans. ASME, Ser. D, Vol. 82, pp. 35～45, (1960).
[148] B. D. O. Anderson: An Algebraic Solution to the Spectral Factrization Problem; IEEE Trans. Automatic Contr., Vol. AC-12, No. 4, pp. 410～414, Aug., 1967.
[149] B. D. O. Anderson: The Inverse Problem of Stationary Covariance Generation; J. Statist. Phys., Vol. 1, pp. 133～147, (1969).
[150] J. C. Willems: Stationary Covariance Generation via the Algebraic Riccati Equation; in Proc. 4 th UKAC Control Convention Multivariable Control System Design and Applications, pp. 24～29, Sept. (1971).
[155] T. Yoshizawa: Stability Theory by Liapunov's Second Method; Math. Soc. of Japan, (1966).
[156] A. Halanay: Differential Equations, Stability, Oscillations, Time Lags; Academic Press, New York, (1966).
[157] R. E. Kalman & J. E. Bertram: Control System Analysis and Design Via the "Second Method" of Lyapunov' I Continuous-Time Systems; ASME. Journal of Basic Engineeg, pp. 371～393, (1960).

[158] J. P. Lasalle & S. Lefschetz: Stability by Liapunov's Direct Method; Academic Press, New York, (1961).
[159] J. K. Hale: Ordinary Differential Equations, John-Wiley, New York, (1969).
[160] R. E. Kalman: Lyapunov Function for the Problem of Lure in Automatic Control; Proc. N. A. S. Vol. 49, pp. 201～205, (1963).
[161] V. A. Yacubovich: The Solution of Certain Matrix Inequalities in Automatic Control Theory; Dokl. Akad. Nauk USSR, 143, pp. 1304～1307, (1962).
[162] K. R. Meyer: On the Existence of Lyapunov Functions for the Problem of Luré; SIAM, Control, Vol. 3, pp. 373～383, (1966).
[163] V. M. Popov: Absolute Stability of Nonlinear Systems of Automatic Control; Avt. I Telemekh., 22, pp. 961～979, (1961).
[164] C. A. Desoer & A. J. Thomasian: A Note On Zero-State Stability of Linear Systems; Proc. 1st Allerton Conference, pp. 50～51, (1963).
[165] 吉田,河田,岩村: 位相解析の基礎; 岩波, p. 172, (1960).
[166] L. M. Silverman & B. D. O. Anderson: Controllability, Observability and Stability of Linear Systems; SIAM J. Control, 6, 1, pp. 121～130, (1968).

多项式矩阵及其应用[1]（基础部分）

引言

设 F 为一个体，$a^{(k)}(k=0, 1, \cdots, l)$ 为 F 的元，当令 s 为变数时[2]，

$$a(s) \triangleq a^{(l)}s^l + a^{(l-1)}s^{(l-1)} + \cdots + a^{(1)}s + a^{(0)} = \sum_{k=0}^{l} a^{(k)}s^k \tag{1}$$

称为变量 s 的多项式[3]．当 $m\times n$ 矩阵 $\boldsymbol{A}$ 的元素 a_{ij} 是 s 的多项式

$$a_{ij}(s) \triangleq a_{ij}^{(l)}s^l + a_{ij}^{(l-1)}s^{l-1} + \cdots + a_{ij}^{(0)},\ i=1, \cdots, m;\ j=1, \cdots, n \tag{2}$$

时，$\boldsymbol{A}(s)$ 称为多项式矩阵[4]．下面我们来介绍这种矩阵的基本性质及其在自动控制中的若干应用．以下设 F 为复数体 C，利用矩阵

$$\boldsymbol{A}_k \triangleq [a_{ij}^{(k)}] \in C^{m\times n},\ k=0, 1, \cdots, l \tag{3}$$

则 $\boldsymbol{A}(s)$ 亦可表示成

$$\boldsymbol{A}(s) = \boldsymbol{A}_l s^l + \boldsymbol{A}_{l-1}s^{l-1} + \cdots + \boldsymbol{A}_1 s + \boldsymbol{A}_0. \tag{4}$$

$\boldsymbol{A}(s)$ 的子式仍然是 s 的多项式．当 $r\times r$ 子矩阵中至少有一个行列式不恒等于零，而 $(r+1)\times(r+1)$ 以上子矩阵的行列式均恒等于零时，则 r 称为 $\boldsymbol{A}(s)$ 的秩[5]．对于以多项式为元素的向量 $\boldsymbol{a}_i(s)\ (i=1, 2, \cdots, r)$，当选择不全为零的多项式 $\alpha_i(s)\ (i=1, 2, \cdots, r)$，能使

$$\sum_{i=1}^{r} \alpha_i(s)\boldsymbol{a}_i(s) = \boldsymbol{0} \tag{5}$$

成立时，称为 $\boldsymbol{a}_i(s)$ 线性相关．当仅限于 $\alpha_i(s)\equiv 0\ (i=1, 2, \cdots, r)$ 的情况下(5)式才成立时，称为 $\boldsymbol{a}_i(s)$ 线性独立．在 $\boldsymbol{A}(s)$ 的秩和列(行)的线性独立性之间有下列关系

rank $\boldsymbol{A}(s)=r \Leftrightarrow \boldsymbol{A}(s)$ 中有 r 个列(行)线性独立，而 $(r+1)$ 个以上列(行)的任何组合均线性相关

若 $n\times n$ 阶方阵 $\boldsymbol{A}(s)$ 的秩为 n，则 $\det \boldsymbol{A}(s)$ 为不恒等于 0 的多项式．这里再使条件严格一些，假定 $\det \boldsymbol{A}(s)$ 是不为 0 的常数．于是

$$[\boldsymbol{A}(s)]^{-1} = \frac{\operatorname{adj}\boldsymbol{A}(s)}{\det \boldsymbol{A}(s)} \tag{6}$$

也是多项式矩阵．这种多项式矩阵称为幺模阵[6]（unimodular matrix）．

基本变换

所谓多项式矩阵的基本行(列)变换，是指下列三种典型操作[7]：

1) 该部分原作为上述讲座的第 III 部分内容，后来适当删减后发表，现分《基础部分》和《应用部分》译出附于本书后面．

2) 多项式中的变量通常多用 λ 表示，为了与下面得出的传递函数对照，这里使用了 s．

3) 设 $k_0 \triangleq \{a^{(k)}\neq 0$ 的最大 $k\}$，则 k_0 称为 $a(s)$ 的次数．以下次数用 $\delta(\cdot)$ 表示，即 $k_0=\delta(a)$．当 $a^{(k_0)}=1$ 时，称为 $a(s)$ 是首一的（参看 P-I-4）．

4) 参看 P-I-4、B-I-12．

5) 参看 B-I-3 中"矩阵的秩"．

6) 参看绪论部分 B-0-1．

7) 参看 B-I-2．

(型式 1)　对于任意一行(列)乘以不为 0 的任意常数;

(型式 2)　在任意行(列)上加上其它行乘以任意多项式;

(型式 3)　两行(列)互相交换.

基本行(列)变换和对 $\boldsymbol{A}(s)$ 左(右)乘以下列 $m\times m$ 矩阵($n\times n$ 矩阵)等价. 亦即, 对于基本行变换.

(型式 1)　对于第 i 行乘以常数 $\alpha\neq 0$: $\boldsymbol{A}_{r1}=\boldsymbol{T}_{r1}\boldsymbol{A}$

$$\boldsymbol{T}_{r1}=\begin{array}{c@{}l}
\left[\begin{array}{ccccccc}
1&&&&&&\mathbf{0}\\
&\ddots&&&&&\\
&&1&&&&\\
&&&\alpha&&&\\
&&&&1&&\\
&&&&&\ddots&\\
\mathbf{0}&&&&&&1
\end{array}\right] & \begin{array}{l}\\ \\ \\ \leftarrow\text{第 } i \text{ 行}\\ \\ \\ \\ \end{array}\\
\begin{array}{ccccccc}&&&\uparrow&&&\\&&&\text{第 } i \text{ 列}&&&\end{array} &
\end{array}\tag{7}$$

(型式 2)　对第 j 行乘以多项式 $\alpha(s)$ 后加到第 i 行: $\boldsymbol{A}_{r2}=\boldsymbol{T}_{r2}\boldsymbol{A}$

$$\boldsymbol{T}_{r2}=\begin{array}{c@{}l}
\left[\begin{array}{cccc}
1&&&\mathbf{0}\\
&\ddots&&\alpha(s)\\
&&\ddots&\\
\mathbf{0}&&&1
\end{array}\right] & \begin{array}{l}\\ \leftarrow\text{第 } i \text{ 行}\\ \\ \\ \end{array}\\
\begin{array}{cccc}&&&\uparrow\\&&&\text{第 } j \text{ 列}\end{array} &
\end{array}\tag{8}$$

(型式 3)　第 i 行和第 j 行互相交换: $\boldsymbol{A}_{r3}=\boldsymbol{T}_{r3}\boldsymbol{A}$

$$\boldsymbol{T}_{r3}=\begin{array}{c@{}l}
\left[\begin{array}{ccccccccccc}
1&&&&&&&&&&\mathbf{0}\\
&\ddots&&&&&&&&&\\
&&1&&&&&&&&\\
&&&0&\cdots&\cdots&\cdots&1&&&\\
&&&\vdots&1&&&\vdots&&&\\
&&&\vdots&&\ddots&&\vdots&&&\\
&&&\vdots&&&1&\vdots&&&\\
&&&1&\cdots&\cdots&\cdots&0&&&\\
&&&&&&&&1&&\\
&&&&&&&&&\ddots&\\
\mathbf{0}&&&&&&&&&&1
\end{array}\right] & \begin{array}{l}\\ \\ \\ \leftarrow\text{第 } i \text{ 行}\\ \\ \\ \\ \leftarrow\text{第 } j \text{ 行}\\ \\ \\ \\ \end{array}\\
\begin{array}{ccccccccccc}&&&\uparrow&&&&\uparrow&&&\\&&&\text{第 } i \text{ 列}&&&&\text{第 } j \text{ 列}&&&\end{array} &
\end{array}\tag{9}$$

这些基本行变换矩阵均为幺模阵, $\boldsymbol{T}_{ri}$ $(i=1,\ 2,\ 3)$ 的逆矩阵是同样型式的基本行变换矩阵. 对于基本列变换,亦有同样结果成立. 而且

[引理 1]　关于多项式方阵 $\boldsymbol{A}(s)$, 下列条件互相等价. (i) $\boldsymbol{A}(s)$ 是幺模阵. (ii) $\boldsymbol{A}(s)$ 可以表示成有限个基本行变换矩阵之积. (iii) $\boldsymbol{A}(s)$ 可以表示成有限个基本列变换矩阵之积.

该引理利用下述定理很容易证明, 故此省略[1]. 利用该结果, 显然, 多项式矩阵 $\boldsymbol{B}(s)=n\times n$ 可以由对 $\boldsymbol{A}(s)=m\times n$ 施行有限次基本变换得到, 与存在着两个幺模阵 $\boldsymbol{P}(s)=m\times m$, $\boldsymbol{Q}(s)=n\times n$, 可以将 $\boldsymbol{B}(s)$ 表示成

$$\boldsymbol{B}(s)=\boldsymbol{P}(s)\boldsymbol{A}(s)\boldsymbol{Q}(s)\tag{10}$$

等价.

1) 参看文献[1],第 135～136 页.

[定义 1] 当多项式矩阵 $\boldsymbol{A}(s)$, $\boldsymbol{B}(s)$对于两个幺模阵$\boldsymbol{P}(s)$, $\boldsymbol{Q}(s)$满足(10)式时,称为 $\boldsymbol{A}(s)$和$\boldsymbol{B}(s)$等价(equivalent).

司密斯典范形[1,2]

下面定理作为以下讨论的基础,是非常重要的.

[定理 1] 任意多项式矩阵 $\boldsymbol{A}(s)$均等价于典范形 $\boldsymbol{A}_c(s)$, 而

$$\boldsymbol{A}_c(s)=\left[\begin{array}{cccc|c} \gamma_1(s) & & & \mathbf{0} & \\ & \gamma_2(s) & & & \mathbf{0} \\ \mathbf{0} & & \ddots & & \\ & & & \gamma_r(s) & \\ \hline & & \mathbf{0} & & \mathbf{0} \end{array}\right]$$

$$r=\operatorname{rank}\boldsymbol{A}(s)\leqslant\min(\mathrm{m},\ \mathrm{n}) \tag{11}$$

式中 $\gamma_1(s)$, $\cdots$, $\gamma_r(s)$ 是不恒为0的首一多项式, 而且 $\gamma_{j+1}(s)$ 可以被 $\gamma_j(s)$ 整除 ($j=1$, 2, $\cdots$, $r-1$).

(证明) 实际上,可以根据求 $\boldsymbol{A}_c(s)$的次序来证明.

① 若 $\boldsymbol{A}(s)\equiv\mathbf{0}$, 则其本身就是典范形. 若 $\boldsymbol{A}(s)\not\equiv\mathbf{0}$, 即至少有一个不恒等于0的元素存在,则

② 在不恒等于0的元素中,将次数最低的元素,利用行交换及列交换换成(1,1)元素.

③ 将第1行、第1列的元素用 $a_{11}(s)$除之,求出其商和余式:

$$a_{i1}(s)=a_{11}(s)q_{i1}(s)+r_{i1}(s),\ i=2,\ 3,\ \cdots,\ m$$
$$a_{1k}(s)=a_{11}(s)q_{1k}(s)+r_{1k}(s),\ k=2,\ 3,\ \cdots,\ n$$

若余式 $r_{i1}(s)$, $r_{1k}(s)$ ($i=2$, $\cdots$, m; $k=2$, $\cdots$, n)均恒等于0, 则可归入下面④. 若余式中至少有一个不恒等于0, 则可以在不恒等于0的余式中找出次数最低的, 例如是 $r_{i_01}(s)$, 将第1行乘以 $q_{i_01}(s)$ 由第 i_0 行减去. 然后回到②. 因余式 $r_{i1}(s)$、$r_{1k}(s)$ 的次数一定比 $a_{11}(s)$ 的次数低,则操作②, ③每进行一次,(1, 1)元素的次数至少降低1次. 因此, 经过有限次之后,所有的余式均恒等于0, 再转入操作④.

④ 由第 i 行减去第1行$\times q_{i1}(s)$ ($i=2$, $\cdots$, m), 而且由第 k 列减去第1列$\times q_{1k}(s)$ ($k=2$, $\cdots$, n). 于是便得到下列形式的多项式矩阵

$$\left[\begin{array}{c|ccc} a_{11}(s) & 0 & \cdots\cdots & 0 \\ \hline 0 & a_{22}(s) & & a_{2n}(s) \\ \vdots & \vdots & & \vdots \\ 0 & a_{m2}(s) & & a_{mn}(s) \end{array}\right] \quad \begin{array}{l}\delta(a_{11})\leqslant\delta(a_{ik}) \\ \forall i,\ k\end{array}$$

若 $a_{ik}(s)$ ($i=2$, $\cdots$, m; $k=2$, $\cdots$, n) 均可被 $a_{11}(s)$整除, 则转入⑤. 若 $a_{ik}(s)$ 中有不能被 $a_{11}(s)$ 整除的, 例如若是 $a_{i_0k_0}(s)$, 则可将第 i_0 行加到第1行,再回到③. 与前面一样,同样可以证明,操作②、③、④经过有限次,必然可转到操作⑤.

⑤ 考察右下角的子矩阵, 对它进行①~④的操作, 因这里右下角的子矩阵的元素均能被 $a_{11}(s)$整除,则对该子矩阵进行基本变换所得到的任意矩阵, 其所有的元素仍然能被 $a_{11}(s)$整除, 注意到这一点, 按照①~⑤的操作, 所处理矩阵的行和列各减少1个, 则经过有限次之后,行或列数将变为0. 这样便转入⑥.

⑥ 在该阶段中，原矩阵 $\boldsymbol{A}(s)$ 变成下列形式

$$\left[\begin{array}{ccc|c} a_{11}(s) & & \boldsymbol{0} & \\ & a_{22}(s) & & \boldsymbol{0} \\ \boldsymbol{0} & & \ddots & \\ & & a_{rr}(s) & \\ \hline & \boldsymbol{0} & & \boldsymbol{0} \end{array}\right] \tag{12}$$

$a_{ii}(s)$ 是不恒为 0 的多项式，如⑤中注意所述，$a_{ii}(s)$ 可整除 $a_{i+1,i+1}(s)$ $(i=1, 2, \cdots, r-1)$. 显然，该矩阵的秩是 r，因进行基本变换不改变矩阵的秩，则 r 等于 rank $\boldsymbol{A}(s)$.

若 $a_{ii}(s)$ 的最高次系数不为 1，用该系数除第 i 行，便将 $a_{ii}(s)$ 变成首一多项式 $r_i(s)$，转入⑦.

⑦ 得到典范形 $\boldsymbol{A}_c(s)$. 证明完毕

$\boldsymbol{A}_c(s)$ 称为司密斯典范形(Smith's Canonical form). $\gamma_j(s)$ 称为 $\boldsymbol{A}(s)$ 的不变因子(invariant polynomial).

[例 1] 试求下列多项式矩阵的司密斯典范形.

$$\boldsymbol{A}(s)=\begin{bmatrix} s^2+9s+8 & 4 & s+3 \\ 0 & s+3 & s+2 \end{bmatrix} \quad \text{将第 1 列和第 2 列交换}$$

$$\to\begin{bmatrix} 4 & s^2+9s+8 & s+3 \\ s+3 & 0 & s+2 \end{bmatrix} \quad \text{由第二行中减去第一行}\times(s+3)/4$$

$$\to\begin{bmatrix} 4 & s^2+9s+8 & s+3 \\ 0 & -0.25(s+3)(s^2+9s+8) & (s+2)-0.25(s+3)^2 \end{bmatrix}$$

由第 2 列减去第 1 列 $\times(s^2+9s+8)/4$,

由第 3 列减去第 1 列 $\times(s+3)/4$

$$\to\begin{bmatrix} 4 & 0 & 0 \\ 0 & -0.25(s+3)(s^2+9s+8) & (s+2)-0.25(s+3)^2 \end{bmatrix} \quad \text{将第 2, 3 列交换}$$

$$\to\begin{bmatrix} 4 & 0 & 0 \\ 0 & -0.25(s+1)^2 & -0.25(s+1)(s+3)(s+8) \end{bmatrix}$$

由第 3 列减去第 2 列 $\times(s+10)$

$$\to\begin{bmatrix} 4 & 0 & 0 \\ 0 & -0.25(s+1)^2 & -14(s+1) \end{bmatrix} \quad \text{将第 2, 3 列交换}$$

$$\to\begin{bmatrix} 4 & 0 & 0 \\ 0 & -14(s+1) & -0.25(s+1)^2 \end{bmatrix} \quad \text{由第 3 列减去第 2 列}\times(s+1)/56$$

$$\to\begin{bmatrix} 4 & 0 & 0 \\ 0 & -14(s+1) & 0 \end{bmatrix} \quad \begin{array}{l}\text{第 1 行}\times 1/4 \\ \text{第 2 行}\times(-1/14)\end{array}$$

$$\to\begin{bmatrix} 1 & 0 & 0 \\ 0 & s+1 & 0 \end{bmatrix} \quad \text{司密斯典范形}$$

不变因子的性质

根据前节说明，所谓以多项式为元素的矩阵 $\boldsymbol{A}(s)=m\times n$ 的不变因子 $\gamma_j(s)$，是 $\boldsymbol{A}(s)$

的司密斯典范形的(j, j)元素$(j=1, \cdots, r)$. 本节先讨论一下更直接地求不变因子的方法. 为此还需要讲点预备知识.

[定义 2] "设矩阵 $\boldsymbol{A}(s)=m\times n$ 的秩为 r. 在能够整除 $\boldsymbol{A}(s)$ 的所有 k 级子式的多项式中取出次数最高的, 即为最大公因式. 其中, 首一多项式是唯一确定的, 设为 $d_k(s)$, $k=0, 1, 2, \cdots, r$. 这些首一最大公因式称为 $\boldsymbol{A}(s)$ 的行列式因子."

[引理 2] "等价矩阵 $\boldsymbol{A}(s)$ 和 $\boldsymbol{B}(s)$ 的行列式因子一致."

(证明) 一般, 在 $\boldsymbol{C}=\boldsymbol{AB}$ 的矩阵 $\boldsymbol{A}\in\widetilde{R}^{m\times p}$, $\boldsymbol{B}\in\widetilde{R}^{p\times n}$, $\boldsymbol{C}\in\widetilde{R}^{m\times n}$ ($\widetilde{R}$ 为可换环) 的 k 级子式之间, 有下列关系

$$\det\boldsymbol{C}\begin{pmatrix} i_1, & i_2, & \cdots, & i_k \\ j_1, & j_2, & \cdots, & j_k \end{pmatrix}$$

$$=\sum_{1\leqslant l<\cdots<l_k\leqslant p}\det\boldsymbol{A}\begin{pmatrix} i_1, & i_2, & \cdots, & i_k \\ l_1, & l_2, & \cdots, & l_k \end{pmatrix}\det\boldsymbol{B}\begin{pmatrix} l_1, & l_2, & \cdots, & l_k \\ j_1, & j_2, & \cdots, & j_k \end{pmatrix}$$

(参看 P-I-2). 将该公式用于等价关系矩阵

$$\boldsymbol{B}(s)=\boldsymbol{P}(s)\boldsymbol{A}(s)\boldsymbol{Q}(s)$$

$$\boldsymbol{A}(s)=\boldsymbol{P}(s)^{-1}\boldsymbol{B}(s)\boldsymbol{Q}(s)^{-1},$$

则 $\boldsymbol{A}(s)$ ($\boldsymbol{B}(s)$) 的 k 级子式可以表示成 $\boldsymbol{B}(s)$ ($\boldsymbol{A}(s)$) 的 k 级子式的线性组合(系数是 s 的多项式). 由此可见, $\boldsymbol{A}(s)$ 和 $\boldsymbol{B}(s)$ 的秩一致. 而且, 若令 $\boldsymbol{A}(s)$ 的 k 级行列式因子为 $d_k(s)$, $\boldsymbol{B}(s)$ 的 k 级行列式因子为 $\tilde{d}_k$, 则 $d_k(s)$ 和 $\tilde{d}_k(s)$ 必能相互整除. 因此, 可得结论 $d_k(s)=\tilde{d}_k(s)$, $(k=0, 1, \cdots, r)$.

证明完毕

将上述结果用于 $\boldsymbol{A}(s)$ 的司密斯典范形, 如下所述, 不变因子可直接由 $\boldsymbol{A}(s)$ 的行列式因子求出. 由此可见, 与变换成司密斯典范形的中间计算过程无关, 它是不变量.

[定理 2]

"设 $\boldsymbol{A}(s)=m\times n$ 的行列式因子为 $d_0(s), d_1(s), \cdots, d_r(s)$, 不变因子为 $\gamma_1(s), \gamma_2(s), \cdots, \gamma_r(s)$, 则

$$\gamma_1(s)=\frac{d_1(s)}{d_0(s)},\ \gamma_2(s)=\frac{d_2(s)}{d_1(s)},\ \cdots,\ \gamma_r(s)=\frac{d_r(s)}{d_{r-1}(s)} \tag{13}$$

成立."

[问题 1] 试对 $\boldsymbol{A}(s)$ 的司密斯典范形用引理 2 证明(13)式.

[问题 2] 试根据行列式因子求出例 1 中矩阵 $\boldsymbol{A}(s)$ 的司密斯典范形.

[例 2] 对于

$$\boldsymbol{J}=\begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ & & & \vdots & \\ 0 & 0 & 0 & \cdots & \lambda \end{bmatrix}\in C^{n\times n}\ (C\text{ 是复数体})$$

$[s\boldsymbol{I}_n-\boldsymbol{J}]$ 的行列式因子是 $d_0=d_1=\cdots=d_{n-1}=1$, $d_n=(s-\lambda)^n$. 因此, 其不变因子为 $\gamma_1=\gamma_2=\cdots=\gamma_{n-1}=1$, $\gamma_n=(s-\lambda)^n$.

[例 3]

$$\boldsymbol{A}_c=\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix},$$

$$\boldsymbol{A}_0=\begin{bmatrix} 0 & 0 & \cdots & -a_0 \\ 1 & 0 & \cdots & -a_1 \\ 0 & 1 & \cdots & -a_2 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & -a_{n-1} \end{bmatrix}(=\boldsymbol{A}_c^T),$$

对于 $\boldsymbol{A}_c\in R^{n\times n}$, $\boldsymbol{A}_0\in R^{n\times n}$, $[s\boldsymbol{I}-\boldsymbol{A}_c]$, $[s\boldsymbol{I}-\boldsymbol{A}_0]$的行列式因子均为 $d_0=d_1=\cdots=d_{n-1}=1$, $d_n=\det[s\boldsymbol{I}-\boldsymbol{A}_c]=\det[s\boldsymbol{I}-\boldsymbol{A}_0]=s^n+a_{n-1}s^{n-1}+\cdots+a_0$. 因此，不变因子也是 $\gamma_1=\gamma_2=\cdots=\gamma_{n-1}=1$, $\gamma_n=d_n$.

以上讨论了一般的以多项式为元素的多项式矩阵. 作为特殊情况，我们来考虑

$$\boldsymbol{A}(s)=[s\boldsymbol{I}_n-\boldsymbol{A}],\ \boldsymbol{A}\in C^{n\times n}\quad (\boldsymbol{I}_n=n\times n\ \text{单位矩阵}).$$

若将 $\boldsymbol{A}(s)$ 限定为这种形式，则可得出对于一般多项式矩阵 $\boldsymbol{A}(s)$ 不成立的各种附加的性质. 首先，因 $[s\boldsymbol{I}_n-\boldsymbol{A}]$ 的秩是 n, 则具有 n 个不变因子. 设这些不变因子为 $\gamma_1(s)$, $\gamma_2(s)$, $\cdots$, $\gamma_n(s)$, 则

$$\det[s\boldsymbol{I}_n-\boldsymbol{A}]=\gamma_1(s)\gamma_2(s)\cdots\gamma_n(s).$$

由(13)式得

$$\gamma_n(s)=\frac{\det[s\boldsymbol{I}_n-\boldsymbol{A}]}{d_{n-1}(s)}$$

$$d_{n-1}(s)=[s\boldsymbol{I}_n-\boldsymbol{A}]\text{的}(n-1)\text{级子式的最大公因式}$$

因 $d_{n-1}(s)$ 也是 $\operatorname{adj}[s\boldsymbol{I}_n-\boldsymbol{A}]$ 的元素的最大公因式，则 $\gamma_n(s)$ 也可以说是由 $\boldsymbol{A}$ 的特征多项式中去掉 $\operatorname{adj}[s\boldsymbol{I}_n-\boldsymbol{A}]$ 的元素的最大公因式后剩下的多项式. 由此可见[2]，多项式 $\gamma_n(s)$ 是 $\boldsymbol{A}$ 的最小多项式.

对于[例 3]中的 $\boldsymbol{A}_0$, $\boldsymbol{A}_c$, 均为 $d_{n-1}(s)=1$, 则其最小多项式等于特征多项式.

对于一般的多项式元素矩阵 $\boldsymbol{A}(s)=m\times n$, $\boldsymbol{B}(s)=m\times n$, 根据[定理 1]，显然下列关系成立:

$$\boldsymbol{A}(s)\text{和}\boldsymbol{B}(s)\text{等价}\Leftrightarrow\boldsymbol{A}(s)\text{和}\boldsymbol{B}(s)\text{的不变因子一致}\tag{14}$$

对于$[s\boldsymbol{I}-\boldsymbol{A}]$, $[s\boldsymbol{I}-\boldsymbol{B}]$形的多项式矩阵，有比这个更强的关系[1,2]，即

$$\boldsymbol{A}\text{ 和 }\boldsymbol{B}\text{ 相似(即},\boldsymbol{B}=\boldsymbol{PAP}^{-1},\ \det\boldsymbol{P}\neq0)\Leftrightarrow[s\boldsymbol{I}-\boldsymbol{A}]\text{ 和 }[s\boldsymbol{I}-\boldsymbol{B}]\text{的不变因子一致}\tag{15}$$

若 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 相似，则$[s\boldsymbol{I}-\boldsymbol{A}]$和$[s\boldsymbol{I}-\boldsymbol{B}]$必然等价. 因此，$[s\boldsymbol{I}-\boldsymbol{A}]$和$[s\boldsymbol{I}-\boldsymbol{B}]$的不变因子一致. (15)式意味着，反过来，若$[s\boldsymbol{I}-\boldsymbol{A}]$和$[s\boldsymbol{I}-\boldsymbol{B}]$的不变因子相等，则 $\boldsymbol{P}(s)=\boldsymbol{P}$, $\boldsymbol{Q}(s)=\boldsymbol{P}^{-1}$. $[s\boldsymbol{I}-\boldsymbol{B}]=\boldsymbol{P}(s)[s\boldsymbol{I}-\boldsymbol{A}]\boldsymbol{Q}(s)$, $\boldsymbol{P}(s)=n\times n$, $\boldsymbol{Q}(s)=n\times n$的等价关系成立.

利用(15)式可以很清楚地找出矩阵 $\boldsymbol{A}\in C^{n\times n}$ 的若当标准形与其不变因子之间的关系. 现在考虑下列若当标准形 $\boldsymbol{J}$

$$
\boldsymbol{J}=\left[\begin{array}{cc|ccc|cc}
\lambda_1 & 1 & & & & & \\
0 & \lambda_1 & & & \mathbf{0} & & \\
\hline
 & & \lambda_1 & 1 & 0 & & \\
 & & 0 & \lambda_1 & 1 & & \\
 & & 0 & 0 & \lambda_1 & & \\
\hline
 & & & & & \lambda_2 & 1 \\
 & \mathbf{0} & & & & 0 & \lambda_2
\end{array}\right].
$$

若将[例 2]用于 $\boldsymbol{J}$ 的各主对角线的子块,则 $(s\boldsymbol{I}-\boldsymbol{J})$ 变换成

$$
\left[\begin{array}{cc|ccc|cc}
1 & 0 & & & & & \\
0 & (s-\lambda_1)^2 & & & & \mathbf{0} & \\
\hline
 & & 1 & 0 & 0 & & \\
 & & 0 & 1 & 0 & & \\
 & & 0 & 0 & (s-\lambda_1)^3 & & \\
\hline
 & & & & & 1 & 0 \\
 & \mathbf{0} & & & & 0 & (s-\lambda_2)^2
\end{array}\right].
$$

再根据行和列的交换,得

$$
\begin{bmatrix}
1 & & & & & & \\
 & 1 & & & & & \\
 & & 1 & & & & \\
 & & & 1 & & & \\
 & & & & (s-\lambda_1)^2 & & \\
 & & & & & (s-\lambda_1)^3 & \\
 & & & & & & (s-\lambda_2)^2
\end{bmatrix}
$$

其司密斯典范形为

$$
\begin{bmatrix}
1 & & & & & & \\
 & 1 & & & & & \\
 & & 1 & & & & \\
 & & & 1 & & & \\
 & & & & 1 & & \\
 & & & & & (s-\lambda_1)^2 & \\
 & & & & & & (s-\lambda_1)^3(s-\lambda_2)^2
\end{bmatrix},
$$

即 $[s\boldsymbol{I}-\boldsymbol{J}]$ 的不变因子为 $\gamma_1=\gamma_2=\gamma_3=\gamma_4=\gamma_5=1$, $\gamma_6=(s-\lambda_1)^2$, $\gamma_7=(s-\lambda_1)^3(s-\lambda_2)^2$. 综上所述,也就是说,根据(15)式,具有这些不变因子的矩阵 $\boldsymbol{A}\in C^{n\times n}$ 能相似变换成上述若当标准形 $\boldsymbol{J}$. 对于一般情况,该结论可叙述如下.

[定理 3]

"当给出 $\boldsymbol{A}\in C^{n\times n}$ 时,设 $[s\boldsymbol{I}-\boldsymbol{A}]$ 的不变因子为 $\gamma_1(s)$, $\gamma_2(s)$, $\cdots$, $\gamma_n(s)$. 这些不变因子作为一次因子的积表示如下:

$$
\begin{aligned}
\gamma_1(s)&=(s-\lambda_1)^{n_{11}}(s-\lambda_2)^{n_{21}}\cdots(s-\lambda_\sigma)^{n_{\sigma 1}},\ \lambda_i\neq\lambda_j,\ i\neq j\\
\gamma_2(s)&=(s-\lambda_1)^{n_{12}}(s-\lambda_2)^{n_{22}}\cdots(s-\lambda_\sigma)^{n_{\sigma 2}},\ \lambda_i\in C\,(i=1,\ 2,\ \cdots,\ \sigma)\\
&\ \ \vdots\\
\gamma_n(s)&=(s-\lambda_1)^{n_{1n}}(s-\lambda_2)^{n_{2n}}\cdots(s-\lambda_\sigma)^{n_{\sigma n}}
\end{aligned}
\tag{16}
$$

因其中 $\gamma_k|\gamma_{k+1}(k=1,\ 2,\ \cdots,\ n)$[1], 则在(16)式的指数之间, $0\leqslant n_{j1}\leqslant n_{j2}\leqslant\cdots\leqslant n_{jn}(j=1,$

1) $\gamma_k|\gamma_{k+1}$ 表示多项式 γ_k 能整除 γ_{k+1}.

2, …, σ)成立．仅将不为0的取出，令

$$0=n_{j1}=n_{j2}=\cdots=n_{j\alpha_j-1}<n_{j\alpha_j}\leqslant\cdots\leqslant n_{jn}\quad(j=1,\ 2,\ \cdots,\ \sigma)$$

这时，$\boldsymbol{A}$ 相似于下列若当标准形

$$\boldsymbol{J}=\begin{bmatrix}\boldsymbol{J}_1 & & & \\ & \boldsymbol{J}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{J}_\sigma\end{bmatrix}\in C^{n\times n},$$

$$\boldsymbol{J}_j=\begin{bmatrix}\boldsymbol{J}_{j\alpha_j} & & & \\ & \boldsymbol{J}_{j\alpha_{j+1}} & & \\ & & \ddots & \\ & & & \boldsymbol{J}_{jn}\end{bmatrix}\in C^{m_j\times m_j},\ \sum_{j=1}^{\sigma}m_j=n,\ j=1,\ 2,\ \cdots,\ \sigma$$

$$\boldsymbol{J}_{jk}=\begin{bmatrix}\lambda_j & 1 & 0\cdots & 0\\ 0 & \lambda_j & 1 & \vdots\\ \vdots & & \ddots & \vdots\\ \vdots & & \ddots & 1\\ 0 & 0 & 0\cdots & \lambda_j\end{bmatrix}\in C^{n_{jk}\times n_{jk}},\ \sum_{k=\alpha_j}^{n}n_{jk}=m_j,\ k=\alpha_j,\ \alpha_{j+1},\ \cdots,\ n\text{”}$$

与 $\boldsymbol{A}$ 的特征值 λ_j 对应的分块矩阵 $\boldsymbol{J}_j$ 中，最大的对角子块是 $\boldsymbol{J}_{jn}$，其大小与包含在 $\gamma_n(s)$ 中的 $(s-\lambda_j)$ 的幂 n_{jn} 相对应，是 $n_{jn}\times n_{jn}$．因此，当且仅当 $n_{jn}=1(j=1,\ 2,\ \cdots)$ 时，$\boldsymbol{J}$ 才能变成对角形矩阵．

亦即，$\boldsymbol{A}$ 能用相似变换化成对角形的充分和必要条件是，不变因子 $\gamma_n(s)$（=最小多项式）能分解成相异的一次因子的乘积．

互素矩阵[3]

在多项式矩阵 $\boldsymbol{A}(s)=m\times n$, $\boldsymbol{B}(s)=m\times p$, $\boldsymbol{C}(s)=p\times n$ 之间，当关系式

$$\boldsymbol{A}(s)=\boldsymbol{B}(s)\boldsymbol{C}(s)$$

成立时，$\boldsymbol{B}(s)$ 称为 $\boldsymbol{A}(s)$ 的左因子(left divisor)(而且 $\boldsymbol{A}(s)$ 称为 $\boldsymbol{B}(s)$ 的右倍式(right multiple))．当 $\boldsymbol{B}(s)$ 分别为 $\boldsymbol{A}_1(s)$, $\boldsymbol{A}_2(s)$ 的左因子时，则称为 $\boldsymbol{B}(s)$ 是 $\boldsymbol{A}_1(s)$, $\boldsymbol{A}_2(s)$ 的左公因子．对于 $\boldsymbol{A}_1(s)$, $\boldsymbol{A}_2(s)$，当 $\boldsymbol{D}(s)$ 是左公因子，而且任意左公因子均能变成 $\boldsymbol{D}(s)$ 的左因子时，即对于任意左公因子 $\boldsymbol{B}(s)$，有满足

$$\boldsymbol{D}(s)=\boldsymbol{B}(s)\boldsymbol{X}(s)$$

的多项式矩阵 $\boldsymbol{X}(s)$ 存在时，称 $\boldsymbol{D}(s)$ 是 $\boldsymbol{A}_1(s)$ 和 $\boldsymbol{A}_2(s)$ 的最大左公因子(greatest common left divisor)，用英文缩写字 gcld 表示．

对于任意 $\boldsymbol{A}_1(s)=m\times n_1$, $\boldsymbol{A}_2(s)=m\times n_2$, gcld 存在．实际上，$[\boldsymbol{A}_1(s)\boldsymbol{A}_2(s)]$ 就是 gcld(中的一个)．

以下当谈到 gcld，皆指其中的 $m\times m$ 方阵．

[引理 3]　对于任意 $\boldsymbol{A}_1(s)=m\times n_1$, $\boldsymbol{A}_2(s)=m\times n_2$,

(a) 存在 $m\times m$ 方阵 gcld.

(b) 对于 $m\times m$ gcld $\boldsymbol{G}(s)$，有满足

$$\boldsymbol{A}_1(s)\boldsymbol{R}_1(s)+\boldsymbol{A}_2(s)\boldsymbol{R}_2(s)=\boldsymbol{G}(s)$$

的 $\boldsymbol{R}_1(s)=n_1\times m$, $\boldsymbol{R}_2(s)=n_2\times m$ 存在．

(c) 设 $\boldsymbol{G}_1(s)=m\times m$ 为其中的一个 gcld, $\boldsymbol{T}(s)=m\times m$ 为任意幺模阵时, 则 $\boldsymbol{G}_2(s)\triangleq\boldsymbol{G}_1(s)\boldsymbol{T}(s)$ 仍然是 gcld.

(证明) (a) 设合成矩阵

$$[\underbrace{\boldsymbol{A}_1(s)}_{n_1}\underbrace{\boldsymbol{A}_2(s)}_{n_2}\underbrace{\boldsymbol{0}}_{\nu}]\}m,\ \nu\triangleq\max\{0,\ m-n_1-n_2\}$$

的司密斯典范形为 $\boldsymbol{\Gamma}(s)$, 则下列关系式成立:

$$[\boldsymbol{A}_1(s)\boldsymbol{A}_2(s)\boldsymbol{0}]=\boldsymbol{P}(s)\boldsymbol{\Gamma}(s)\boldsymbol{Q}(s) \tag{17}$$

$\boldsymbol{P}(s)=m\times m,\ \boldsymbol{Q}(s)=(n_1+n_2+\nu)\times(n_1+n_2+\nu)$ 均为幺模阵

$$\boldsymbol{\Gamma}(s)=[\underbrace{\boldsymbol{\Gamma}_1(s)}_{m}\underbrace{\boldsymbol{0}}_{\bar{\nu}}]\}m, \tag{18}$$

$$\bar{\nu}\triangleq\max\{0,\ n_1+n_2-m\}$$

$$\boldsymbol{\Gamma}_1(s)=\operatorname{diag}\{\gamma_1(s),\ \cdots,\ \gamma_r(s),\ 0,\ \cdots,\ 0\}$$

$$r\triangleq\operatorname{rank}[\boldsymbol{A}_1(s)\boldsymbol{A}_2(s)].$$

现将 $\boldsymbol{Q}(s)$, $\boldsymbol{Q}^{-1}(s)$ 分别划分成相应的分块矩阵, 设

$$\boldsymbol{Q}(s)=\begin{array}{c}\left[\begin{array}{c:c:c}\boldsymbol{Q}_{11}(s)&\boldsymbol{Q}_{12}(s)&\boldsymbol{Q}_{13}(s)\\ \hdashline \boldsymbol{Q}_{21}(s)&\boldsymbol{Q}_{22}(s)&\boldsymbol{Q}_{23}(s)\end{array}\right]\begin{array}{l}\}m\\ \}\bar{\nu}\end{array}\\ \begin{array}{ccc}\quad n_1\quad&\quad n_2\quad&\quad\nu\quad\end{array}\end{array} \tag{19.1}$$

$$\boldsymbol{Q}^{-1}(s)=\begin{array}{c}\left[\begin{array}{c:c}\boldsymbol{R}_{11}(s)&\boldsymbol{R}_{12}(s)\\ \hdashline \boldsymbol{R}_{21}(s)&\boldsymbol{R}_{22}(s)\\ \hdashline \boldsymbol{R}_{31}(s)&\boldsymbol{R}_{32}(s)\end{array}\right]\begin{array}{l}\}n_1\\ \}n_2,\\ \}\nu\end{array}\\ \begin{array}{cc}\quad m\quad&\quad\bar{\nu}\quad\end{array}\end{array} \tag{19.2}$$

根据(17), (18), (19.1)式,

$$\boldsymbol{A}_i(s)=\boldsymbol{P}(s)\boldsymbol{\Gamma}_1(s)\boldsymbol{Q}_{1i}(s),\ i=1,\ 2$$

即 $\boldsymbol{G}_0(s)\triangleq\boldsymbol{P}(s)\boldsymbol{\Gamma}_1(s)=m\times m$ 是 $\boldsymbol{A}_1(s)$, $\boldsymbol{A}_2(s)$ 的左公因子. 其次, 在(17)式两边右乘以 $\boldsymbol{Q}^{-1}(s)$, 利用(19.2)式得

$$[\boldsymbol{A}_1(s)\boldsymbol{A}_2(s)]\begin{bmatrix}\boldsymbol{R}_{11}(s)\\ \boldsymbol{R}_{21}(s)\end{bmatrix}=\boldsymbol{P}(s)\boldsymbol{\Gamma}_1(s)=\boldsymbol{G}_0(s) \tag{20}$$

因此, $\boldsymbol{A}_1(s)$, $\boldsymbol{A}_2(s)$ 的左公因子也是 $\boldsymbol{G}_0(s)$ 的左因子, 则 $\boldsymbol{G}_0(s)$ 是 $\boldsymbol{A}_1(s)$, $\boldsymbol{A}_2(s)$ 的一个 gcld.

(b) 由(20)式关于 $\boldsymbol{G}_0(s)$ 成立, 对于任意 $m\times m$ gcld $\boldsymbol{G}(s)$, 有满足

$$\boldsymbol{G}(s)=\boldsymbol{G}_0(s)\boldsymbol{X}(s)$$

的 $\boldsymbol{X}(s)$ 存在. 因此

$$\boldsymbol{A}_1(s)\{\boldsymbol{R}_{11}(s)\boldsymbol{X}(s)\}+\boldsymbol{A}_2(s)\{\boldsymbol{R}_{21}(s)\boldsymbol{X}(s)\}=\boldsymbol{G}(s)$$

(c) 设 $\boldsymbol{G}_1(s)$ 是 gcld 中的一个,

$$\boldsymbol{A}_i(s)=\boldsymbol{G}_1(s)\boldsymbol{Y}_i(s),\ i=1,\ 2$$

及 $\boldsymbol{B}(s)$ 为任意左公因子,

$$\boldsymbol{G}_1(s)=\boldsymbol{B}(s)\boldsymbol{X}(s)$$

成立, 所以

$$\boldsymbol{A}_i(s)=\boldsymbol{G}_1(s)\boldsymbol{T}(s)\boldsymbol{T}^{-1}(s)\boldsymbol{Y}_i(s)=\boldsymbol{G}_2(s)\{\boldsymbol{T}^{-1}(s)\boldsymbol{Y}_i(s)\};\ i=1,\ 2$$

$G_2(s)$ 也是左公因子，而且

$$G_2(s)=G_1(s)T(s)=B(s)\{X(s)T(s)\}$$

因此，$G_2(s)$ 是 $m\times m$ 方阵 gcld.

证明完毕

若对所讨论的 $A_1(s)$, $A_2(s)$ 加以限制，则可以得到比以上更多的性质，存在下列有用结果.

［引理 3′］ 设 $A_1(s)=m\times n_1$, $A_2(s)=m\times n_2$, $\mathrm{rank}[A_1(s)A_2(s)]=m$ 1), 在 $m\times m$ 多项式方阵中，设 $A_1(s)$ 和 $A_2(s)$ 的所有 gcld 的集合为 $\mathscr{G}$, 则

(a) $\det G(s)\neq 0$, $^{\forall}G(s)\in\mathscr{G}$

(b) 对于任意 $G_1(s)\in\mathscr{G}$, $G_2(s)\in\mathscr{G}$, 存在幺模阵 $T(s)=m\times m$, 使得

$$G_1(s)=G_2(s)T(s) \tag{21}$$

成立.

(c) 若某个 $G_1(s)\in\mathscr{G}$ 是幺模阵，则任意 $G\in\mathscr{G}$ 是幺模阵.

（证明） (a) 注意到引理 3(a) 证明中定义的 $G_0(s)$, 根据假定 $\mathrm{rank}[A_1(s)A_2(s)]=m$,

$$\Gamma_1(s)=\mathrm{diag}\{\gamma_1(s),\ \cdots,\ \gamma_m(s)\};\ \det\Gamma_1(s)\neq 0 \tag{22}$$

因此 $\det G_0(s)\neq 0$. 因 $G_0(s)$ 是 gcld, 则对于任意 $G_1(s)\in\mathscr{G}$, 有满足

$$G_0(s)=G_1(s)T_1(s),\ G_1(s)=G_0(s)\widetilde{T}_1(s) \tag{23}$$

的多项式矩阵 $T_1(s)$, $\widetilde{T}_1(s)$ 存在. 在(23)式的第一式中，因 $\det G_0(s)\neq 0$, 则 $\det G_1(s)\neq 0$, 于是(a)得证. 此外，$\det T_1(s)\neq 0$, 故由第 2 式得 $\det\widetilde{T}\neq 0$. 将第 2 式代入第 1 式，得

$$G_0(s)=G_0(s)\widetilde{T}_1(s)T_1(s),$$

因 $\det G_0(s)\neq 0$, 则

$$\widetilde{T}_1(s)T_1(s)=I_m.$$

这意味着 $T_1(s)$, $\widetilde{T}_1(s)$ 皆为幺模阵. 对于其它任意 $G_2(s)\in\mathscr{G}$, 同样也有下列关系

$$G_0(s)=G_2(s)T_2(s),\ G_2(s)=G_0(s)\widetilde{T}_2(s)$$

$$T_2(s),\ \widetilde{T}_2(s):\text{幺模阵}$$

因而

$$G_1(s)=G_2(s)T_2(s)\widetilde{T}_1(s)$$

$$\{T_2(s)\widetilde{T}_1(s)\}:\text{幺模阵}$$

则(21)式得证.

(c) 由(b)之(21)式显然.

证明完毕

（注意）

(i) 引理 3 表示，若求出$[A_1(s)A_2(s)]$的司密斯典范形，则立即可以得到 gcld. 麦克达菲[3](MacDuffee)给出更简单一些的求 gcld 的算法，其基本思想与此相同.

(ii) 引理 3′(b)表示，根据某个 $m\times m$ gcld 阵，其它所有的 $m\times m$ gcld 均可用乘以幺模阵构成. 若不假定 $\mathrm{rank}[A_1(s)A_2(s)]=m$, 则该结果不成立. 即若 $\mathrm{rank}[A_1(s)A_2(s)]$

1) 例如，若 $n=m$, $\det A_1(s)\neq 0$, 则该条件满足.

$<m$，则对于任意两个 $m\times m$ gcld $\boldsymbol{G}_1(s)$，$\boldsymbol{G}_2(s)$，不一定存在满足 $\boldsymbol{G}_2(s)=\boldsymbol{G}_1(s)\boldsymbol{T}(s)$ 的幺模阵 $\boldsymbol{T}(s)=m\times m$.

［定义 3］ 多项式矩阵 $\boldsymbol{A}_1(s)=m\times n_1$，$\boldsymbol{A}_2(s)=m\times n_2$，$(n_1+n_2\geqslant m)$. 当 $[\boldsymbol{A}_1(s)\boldsymbol{A}_2(s)]$ 的司密司典范形是 $[\boldsymbol{I}_m\boldsymbol{0}]$ 时，称为 $\boldsymbol{A}_1(s)$，$\boldsymbol{A}_2(s)$ 相互左素(left coprime).

［定理 4］ 对于多项式矩阵 $\boldsymbol{A}_1(s)=m\times n_1$，$\boldsymbol{A}_2(s)=m\times n_2$，$\operatorname{rank}[\boldsymbol{A}_1(s)\boldsymbol{A}_2(s)]=m$，下列三个条件相互等价.

(a) $\boldsymbol{A}_1(s)$ 和 $\boldsymbol{A}_2(s)$ 相互左素.

(b) $\boldsymbol{A}_1(s)$ 和 $\boldsymbol{A}_2(s)$ 的 $m\times m$ gcld 是幺模阵.

(c) 有满足 $\boldsymbol{A}_1(s)\boldsymbol{B}_1(s)+\boldsymbol{A}_2(s)\boldsymbol{B}_2(s)=\boldsymbol{I}_m$ 的多项式矩阵 $\boldsymbol{B}_1(s)=n_1\times m$，$\boldsymbol{B}_2(s)=n_2\times m$ 存在.

（证明） (a)$\Rightarrow$根据引理 3 和 $\boldsymbol{\Gamma}_1(s)=\boldsymbol{I}_m$，$\boldsymbol{P}(s)$ 是 gcld. 因 $\boldsymbol{P}(s)$ 是幺模阵，由引理 3′，则所有的 $m\times m$ gcld 是幺模阵$\Rightarrow$(b)$\Rightarrow$因 $\boldsymbol{P}(s)\boldsymbol{\Gamma}_1(s)$ 是 gcld，则为幺模阵. 因此，在(20)式两边右乘以 $[\boldsymbol{P}(s)\boldsymbol{\Gamma}_1(s)]^{-1}$，定义 $\boldsymbol{B}_i(s)\triangleq\boldsymbol{R}_{i1}(s)[\boldsymbol{P}(s)\boldsymbol{\Gamma}_1(s)]^{-1}$，$i=1, 2\Rightarrow$(c)$\Rightarrow$ $\operatorname{rank}\boldsymbol{\Gamma}_1(s)=\operatorname{rank}[\boldsymbol{A}_1(s)\boldsymbol{A}_2(s)]=m$，$^\forall s$. 若 $\gamma_m(s)$ 为多项式，则对于满足 $\gamma_m(s_j)=0$ 的 s_j，得出 $\operatorname{rank}\boldsymbol{\Gamma}_1(s_j)<m$，矛盾$\Rightarrow\gamma_1(s)=\cdots=\gamma_m(s)=1\Rightarrow$(a)

证明完毕

（注意） (i) 以上所讲到的关于“左”的定义和性质，适当改述一下对于“右”亦均成立. 为了慎重起见，将定理 4 对于“右”叙述如下.

［定理 4′］ 设

$$\boldsymbol{A}_1(s)=m_1\times n,\ \boldsymbol{A}_2(s)=m_2\times n,\ \operatorname{rank}\begin{bmatrix}\boldsymbol{A}_1(s)\\\boldsymbol{A}_2(s)\end{bmatrix}=n,$$

则下面三个性质相互等价.

(a) 合成矩阵 $\begin{bmatrix}\boldsymbol{A}_1(s)\\\boldsymbol{A}_2(s)\end{bmatrix}$ 的司密斯典范形是 $\begin{bmatrix}\boldsymbol{I}_m\\\boldsymbol{0}\end{bmatrix}$（即 $\boldsymbol{A}_1(s)$，$\boldsymbol{A}_2(s)$ 相互右素）.

(b) $\boldsymbol{A}_1(s)$ 和 $\boldsymbol{A}_2(s)$ 的 $n\times n$ 最大右公因子(greatest common right divisor；用 gcrd 表示)是幺模阵.

(c) 有满足 $\boldsymbol{B}_1(s)\boldsymbol{A}_1(s)+\boldsymbol{B}_2(s)\boldsymbol{A}_2(s)=\boldsymbol{I}_n$ 的多项式矩阵 $\boldsymbol{B}_1(s)=n\times m_1$，$\boldsymbol{B}_2(s)=n\times m_2$ 存在.

(ii) 由上述定理等显而易见，这里关于多项式矩阵所定义的相互左(右)素的性质，可以看成是关于纯量多项式互素性质[1]的推广.

(iii) 两个矩阵相互左素，但未必相互右素. 其逆亦真. 例如，对于

$$\boldsymbol{A}_1(s)=\begin{bmatrix}s^2 & -1\\-s & s^2\end{bmatrix},\ \boldsymbol{A}_2(s)=\begin{bmatrix}s & -s\\0 & 1\end{bmatrix},$$

$\boldsymbol{I}_2$ 是 gcld. 因此，$\boldsymbol{A}_1(s)$，$\boldsymbol{A}_2(s)$ 是左素的. 此外，

$$\boldsymbol{G}_R(s)=\begin{bmatrix}s & 0\\0 & 1\end{bmatrix}$$

是 gcrd，因此 $\boldsymbol{A}_1(s)$，$\boldsymbol{A}_2(s)$ 不是右素的.

1) 参看 P-I-4.

列(行)适宜矩阵

多项式矩阵 $\boldsymbol{A}(s)=n\times n$ 在下列情况下称为是列(行)适宜的: $\det\boldsymbol{A}(s)\neq 0$, 而且当令 $\sigma_j\triangleq$ 构成第 j 列(行)元素的多项式的最高次数(以下称 σ_j 为第 j 列(行)的次数)时

$$\delta(\det\boldsymbol{A}(s))=\sum_{j=1}^{n}\sigma_j \tag{24}$$

成立. 若矩阵

$$\boldsymbol{A}_0=[a_{ij}^0]\in C^{n\times n}$$

按

$$a_{ij}^0\triangleq\boldsymbol{A}(s)\text{的}(i,\ j)\text{元素的}s^{\sigma_j}(s^{\sigma_i})\text{的系数} \tag{25}$$

确定, 显然

$$\boldsymbol{A}(s)\text{列(行)适宜}\Leftrightarrow\det\boldsymbol{A}_0\neq 0 \tag{26}$$

一般, 由 $\delta(d(s))\geqslant\delta(n(s))$ 的两个多项式构成的有理函数 $h(s)=n(s)/d(s)$ 称为是适宜的[1]. 那么, 由多项式矩阵 $\boldsymbol{D}(s)=r\times r$ 和 $\boldsymbol{N}(s)=m\times r$ 构成的有理函数矩阵 $\boldsymbol{H}(s)=\boldsymbol{N}(s)\boldsymbol{D}^{-1}(s)$ 在什么情况下是适宜的(即各元素适宜)?

[引理 5][4,7]

设 $\boldsymbol{D}(s)$ 是列适宜的, 这时

$$\boldsymbol{H}(s)=\boldsymbol{N}(s)\boldsymbol{D}^{-1}(s)\text{是适宜的}\Leftrightarrow\boldsymbol{N}(s)\text{的第 }j\text{ 列次数}\leqslant\boldsymbol{D}(s)\text{的第 }j\text{ 列次数}$$
$$(j=1,\ 2,\ \cdots,\ r)$$

(证明) $\boldsymbol{D}^{-1}(s)=\mathrm{adj}\,\boldsymbol{D}(s)/\det\boldsymbol{D}(s)$. 因 $\boldsymbol{D}(s)$ 是列适宜的, 利用(26)式关系, 则 $\mathrm{adj}\,\boldsymbol{D}(s)$ 变成行适宜的, 其第 j 列的次数是 $n-\sigma_j$, 其中设 $\sigma_j\triangleq\boldsymbol{D}(s)$ 的第 j 行次数, $n\triangleq\sum_{j=1}^{r}\sigma_j$. 因此, 当 $\boldsymbol{N}(s)$ 的第 j 列次数 $\leqslant\sigma_j(j=1,\ 2,\ \cdots,\ r)$ 时, $\boldsymbol{N}(s)\mathrm{adj}\,\boldsymbol{D}(s)$ 的各元素最多为 n 次, 即 $\boldsymbol{H}(s)$ 是适宜的. 反之, 若假定 $\boldsymbol{N}(s)$ 的第 j 列的次数 $>\sigma_j(j=1,\ 2,\ \cdots,\ r)$, 则在 $\boldsymbol{N}(s)\mathrm{adj}\,\boldsymbol{D}(s)$ 的各行中至少有一行, 会出现高于 n 次的元素. 因此, $\boldsymbol{H}(s)$ 不是适宜的.

证明完毕

(注意) 引理 5 是关于列适宜情况下的结果. 一般, 因列适宜矩阵的转置是行适宜矩阵, 故由引理 5 立即可以得到下列关于行适宜情况下的结果:

设 $\boldsymbol{D}(s)=m\times m$ 是行适宜的, 则

$$\boldsymbol{H}(s)=\boldsymbol{D}^{-1}(s)\boldsymbol{N}(s)\text{是适宜的}\Leftrightarrow\boldsymbol{N}(s)\text{的第 }j\text{ 行次数}\leqslant\boldsymbol{D}(s)\text{的第 }j\text{ 行次数}$$
$$(j=1,\ 2,\ \cdots,\ m),$$

一般, 对于 $\det\boldsymbol{A}(s)\neq 0$, 不是列(行)适宜的矩阵 $\boldsymbol{A}(s)$, 利用适当的基本列(行)变换可以变换成列(行)适宜的.

[引理 6][4,7]

"当 $\det\boldsymbol{A}(s)\neq 0$ 的 $n\times n$ 多项式矩阵 $\boldsymbol{A}(s)$ 不是列(行)适宜时, 利用适当的幺模阵 $\boldsymbol{Q}(s)$ ($\boldsymbol{P}(s)$) 可以将 $\hat{\boldsymbol{A}}(s)=\boldsymbol{A}(s)\boldsymbol{Q}(s)$ ($\boldsymbol{P}(s)\boldsymbol{A}(s)$) 变成列(行)适宜的."

(证明)

设 σ_j 为 $\boldsymbol{A}(s)$ 第 j 列的次数, 矩阵 $\boldsymbol{A}_0=[a_{ij}^0]\in C^{n\times n}$ 按(25)式确定, 因 $\boldsymbol{A}(s)$ 不是列适

1) 通常, 当 $\delta(d(s))>\delta(n(s))$ 时, 称 $h(s)=n(s)/d(s)$ 是真分式. ——译者注

宜的，则 $\det \boldsymbol{A}_0=0$. 因此，当 $\boldsymbol{A}_0$ 的第 j 列表示成 $\boldsymbol{d}_j$ 时，存在不全为零的常数 $\alpha_j(j=1, 2, \cdots, n)$，使得

$$\sum_{j=1}^{n}\alpha_j\boldsymbol{d}_j=\boldsymbol{0}$$

成立. 现对于与不为 0 的 α_j 相对应的 $\boldsymbol{A}(s)$ 的列比较其次数，设其中最大值为 σ_{j0}. 因 $\det \boldsymbol{A}(s)\neq 0$，则 $\sigma_{j0}>0$. 其次，将 $\boldsymbol{A}(s)$ 的第 j_0 列乘以 α_{j0}，将 ($\boldsymbol{A}(s)$ 的第 j 列 $\times\alpha_j s^{\sigma_{j0}-\sigma_j}$)，$j=1, 2, \cdots, j_0-1, j_0+1, \cdots, n$ 列顺次加到第 j_0 列. 设利用这种基本列变换构成的矩阵为 $\widetilde{\boldsymbol{A}}(s)$，则除了第 j_0 列外，$\widetilde{\boldsymbol{A}}(s)$ 和 $\boldsymbol{A}(s)$ 相等，其第 j_0 列的次数比 σ_{j0} 低. 将 $\widetilde{\boldsymbol{A}}(s)$ 各列的次数改写，设为 $\widetilde{\sigma}_i$，根据上述，则

$$\sum_{j=1}^{n}\sigma_j>\sum_{j=1}^{n}\widetilde{\sigma}_j \tag{27}$$

成立. 但是，因 (27) 式右边存在固定的下限

$$\sum_{j=1}^{n}\widetilde{\sigma}_j\geqslant\delta(\det\widetilde{\boldsymbol{A}}(s))=\delta(\det\boldsymbol{A}(s))$$

故若反复进行上述列变换，经过有限次之后即可得到 $\hat{\boldsymbol{A}}(s)$，对于其列次数 $\hat{\sigma}_i$，(24) 式成立. 亦即，可以得到列适宜矩阵 $\hat{\boldsymbol{A}}(s)$. 显然，对于不是行适宜的 $\boldsymbol{A}(s)$，进行同样的行变换，也可以变换成行适宜的.

证明完毕

[引理 6] 对于不是列(行)适宜的矩阵给出减小列(行)次数和的基本变换.

对于列(行)适宜矩阵，根据和上面同样的方法，可以增减各列(行)的次数，而使列(行)次数的和保持一定.

[引理 7][5]

设列(行)适宜矩阵 $\boldsymbol{A}(s)=n\times n$ 的列(行)次数为 $\sigma_j(j=1, 2, \cdots, n)$. 对于某两个列(行)，例如第 j, k 列(行)，当 $\sigma_j<\sigma_k$ 时，对 $\boldsymbol{A}(s)$ 进行适当的行及列变换，可以得到列(行)次数 $\hat{\sigma}_i(i=1, 2, \cdots, n)$ 为

$$\begin{cases}\hat{\sigma}_i=\sigma_i,\ i\neq j,\ k\\ \hat{\sigma}_j=\sigma_{j+1},\ \hat{\sigma}_k=\sigma_{k-1}\end{cases}$$

的列(行)适宜矩阵 $\hat{\boldsymbol{A}}(s)$.

(证明)

通过适当的行(列)变换，列(行)适宜矩阵总可以变换成，其主对角线上元素至少比同列(行)其它元素次数高一次的矩阵[1)]. 因此，假定 $\boldsymbol{A}(s)$ 已满足该条件. 而且，不失去一般性，设 $\boldsymbol{A}(s)$ 的对角线元素为首一多项式. 以下，对于 $\boldsymbol{A}(s)$ 是列适宜的情况分步骤说明变换过程. 当 $\boldsymbol{A}(s)$ 为行适宜时亦可同样证明.

第一步

将 ($s\times$ 第 j 行) 加到第 k 行. 该结果除了第 j, k 列外，列次数不变. 第 j 列次数变成 σ_j+1.

第二步

设 (k, k) 元素中 s^{σ_k} 的系数为 β. 将 ($-\beta s^{\sigma_k-\sigma_j-1}\times$ 第 j 列) 加到第 k 列. 结果第 k 列的次数低于 σ_k，设该矩阵为 $\hat{\boldsymbol{A}}(s)$.

1) 例如，设 $\boldsymbol{A}(s)$ 是列适宜的，其列次数为 $\sigma_j(j=1, 2, \cdots, n)$. 若 $\boldsymbol{A}_0$ 按 (25) 式确定，则 $\boldsymbol{A}_0^{-1}\boldsymbol{Q}\boldsymbol{A}(s)$ 的第 j 列主对角线上元素的次数是 σ_j，而且第 j 列的其它元素具有 (σ_j-1) 以下的次数.

第三步

由 $\hat{A}(s)$ 的元素的系数构成下列常数矩阵 $Q\in R^{n\times n}$

$$Q=[q_{il}],\ \begin{cases} q_{il}=\hat{A}(s)\text{ 的 }(i,\ l)\text{ 元素中 } s^{\sigma_l}\text{ 的系数 } l\neq j,\ k \\ q_{ij}=\hat{A}(s)\text{ 的 }(i,\ j)\text{ 元素中 } s^{\sigma_j+1}\text{ 的系数} \\ q_{ik}=\hat{A}(s)\text{ 的 }(i,\ k)\text{ 元素中 } s^{\sigma_k-1}\text{ 的系数} \end{cases}$$

因 $\hat{A}(s)$ 是对 $A(s)$ 进行基本变换得到的，则 $\det\hat{A}(s)=\det A(s)$．而且，根据 Q 的定义，$\det\hat{A}(s)$ 的最高次数 $s^{\sigma_1+\sigma_2+\cdots+\sigma_n}$ 的系数是 $\det Q$．由此 $\det Q=1$，故 Q 是正则的．这意味着，$\hat{A}(s)$ 的第 k 列次数是 σ_k-1，而 $\hat{A}(s)$ 是满足定理条件的列适宜矩阵．

证明完毕

（例）

$$A(s)=\begin{bmatrix} s^2+1 & s^3 \\ s & s^5+1 \end{bmatrix}\xrightarrow{\text{第一步}}\begin{bmatrix} s^2+1 & s^3 \\ s^3+2s & s^5+s^4+1 \end{bmatrix}$$

$$\xrightarrow{\text{第二步}}\begin{bmatrix} s^2+1 & -s^4+s^3-s^2 \\ s^3+2s & s^4-2s^3+1 \end{bmatrix}=\hat{A}(s),\quad Q=\begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}$$

此外，我们再来考虑一下列(行)适宜矩阵 $A(s)$ 与其不变因子的关系．H．H．罗森布罗克(Rosenbrock)曾证明，研究在线性系统中状态反馈对系统结构的影响时，该结果具有非常重要的意义．

首先，作为列(行)适宜矩阵，我们来考察一下主对角线元素具有首一多项式的对角形

$$A(s)=\begin{bmatrix} \Delta_{11}(s) & & & \\ & \Delta_{22}(s) & & \\ & & \ddots & \\ & & & \Delta_{nn}(s) \end{bmatrix}\quad \delta(\Delta_{ii}(s))=\sigma_i\quad (i=1,\ 2,\ \cdots,\ n) \tag{28}$$

式中假定各列(行)的次数

$$\sigma_1\geqslant\sigma_2\geqslant\cdots\geqslant\sigma_n\geqslant 0$$

(必要的话要进行列及行的交换)．若令 $A(s)$ 的不变因子为 $\psi_1(s),\psi_2(s),\ \cdots,\ \psi_n(s)$，根据[定理 2]，因

$$\begin{aligned} \psi_1&=A(s)\text{ 的 1 级子式的最大公因式} \\ \psi_1\psi_2&=A(s)\text{ 的 2 级子式的最大公因式} \\ &\ \vdots \\ \psi_1\psi_2\cdots\psi_n&=A(s)\text{ 的 } n\text{ 级子式的最大公因式}(=\det A(s)) \end{aligned}$$

显然，关于其次数

$$\left.\begin{aligned} &\delta(\psi_1)\leqslant\sigma_n \\ &\delta(\psi_1\psi_2)=\delta(\psi_1)+\delta(\psi_2)\leqslant\sigma_n+\sigma_{n-1} \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &\delta(\psi_1\cdots\psi_{n-1})=\delta(\psi_1)+\cdots+\delta(\psi_{n-1})\leqslant\sigma_n+\sigma_{n-1}+\cdots+\sigma_2 \\ &\delta(\psi_1\cdots\psi_n)=\sigma_n+\sigma_{n-1}+\cdots+\sigma_1 \end{aligned}\right\} \tag{29}$$

成立．而且因为是不变因子，则

$$\psi_i(s)/\psi_{i+1}(s),\ i=1,\ 2,\ \cdots,\ n-1 \tag{30}$$

必须成立．

其次，考虑将多项式元素矩阵 $\boldsymbol{B}(s)=n\times n$ 加到 $\boldsymbol{A}(s)$ 的情况，

$$\boldsymbol{A}(s)=\boldsymbol{A}(s)+\boldsymbol{B}(s)$$

其中假定 $\boldsymbol{B}(s)$ 第 j 列(行)($j=1, 2, \cdots, n$)的次数至多为 σ_j-1. 因此，$\boldsymbol{A}(s)$ 是列(行)适宜的，其列(行)次数和 $\boldsymbol{A}(s)$ 相同.

若令 $\boldsymbol{A}(s)$ 的不变因子为 $\psi_1(s), \psi_2(s), \cdots, \psi_n(s)$，则和 $\boldsymbol{A}(s)$ 的不变因子一样，(29)式及(30)式显然成立. 即(29)，(30)式是 $\boldsymbol{A}(s)$ 的不变因子满足的必要条件.

在这里我们要论证的是这样一个事实，反过来，当先给出满足(29)，(30)式的任意首一多项式 $\psi_1(s), \psi_2(s), \cdots, \psi_n(s)$ 时，总可以构成这样的矩阵 $\boldsymbol{B}$，使得 $\boldsymbol{A}(s)=\boldsymbol{A}(s)+\boldsymbol{B}(s)$ 以这些多项式为不变因子. 现在我们对于 $\boldsymbol{A}(s)$ 是列适宜的情况来证明它，而对于行适宜的情况，考虑其转置即可.

[定理 5][5]

当给出(28)式的 $\boldsymbol{A}(s)$ 和满足(29)，(30)式的首一多项式 $\psi_1(s), \psi_2(s), \cdots, \psi_n(s)$ 时，在 $\boldsymbol{A}(s)$ 上加上适当的 $\boldsymbol{B}(s)$，可形成列(行)适宜矩阵 $\boldsymbol{A}(s)=\boldsymbol{A}(s)+\boldsymbol{B}(s)$，它以 $\psi_1(s), \psi_2(s), \cdots, \psi_n(s)$ 为不变因子.

(证明)

设 $\boldsymbol{S}(s)\triangleq \mathrm{diag}(\psi_1(s), \psi_2(s), \cdots, \psi_n(s))$，则 $\boldsymbol{S}(s)$ 是以 $\psi_1(s), \cdots, \psi_n(s)$ 为不变因子的司密斯典范形. 若

$$\delta(\psi_i)=\sigma_{n-i+1},\ i=1, 2, \cdots, n \tag{31}$$

则作为 $\boldsymbol{B}(s)=\widetilde{\boldsymbol{S}}(s)-\boldsymbol{A}(s)$，对角形 $\boldsymbol{B}(s)$ 可很简单地求出，其中 $\widetilde{\boldsymbol{S}}(s)\triangleq\mathrm{diag}(\psi_n(s), \psi_{n-1}(s), \cdots, \psi_1(s))$. 现假定(31)式不成立. 因此，在 $\boldsymbol{S}(s)$ 的列中找出

$$\delta(\psi_i)=\sigma_{n-i+1},\ 1\leqslant i\leqslant j-1,\ \delta(\psi_j)<\sigma_{n-j+1}$$

的前面 j 列. 因

$$\sum_{i=1}^{n}\delta(\psi_i)=\sigma_n+\sigma_{n-1}+\cdots+\sigma_1$$

则对于某个 $k>j$

$$\delta(\psi_j)<\sigma_{n-j+1}<\sigma_{n-k+1}<\delta(\psi_k) \tag{32}$$

成立. 将[引理 7]用于 $\boldsymbol{S}(s)$ 的第 j, k 列，可以使其它列的次数不变，仅使第 j 列的次数增加 1，第 k 列的次数同时减少. 而且，在变换成的 $\widetilde{\boldsymbol{S}}(s)$ 中，主对角线上各元素成为比其同列其它元素次数至少高 1 次的首一多项式.

反复进行这种变换，直到 $\delta(\psi_i)=\sigma_{n-j+1}$ 或 $\delta(\psi_k)=\sigma_{n-k+1}$，或二者同时成立. 若 $\delta(\psi_j)=\sigma_{n-j+1}$，$\sigma_{n-k+1}<\delta(\psi_k)$，则在第 $(j+1)$ 列以右的列中找出 $\delta(\psi_l)<\sigma_{n-l+1}$ 的第 l 列，对第 l 列和第 k 列进行同样的操作. 因为若 $\delta(\psi_j)<\sigma_{n-j+1}$，$\delta(\psi_k)=\sigma_{n-k+1}$，则有 $\delta(\psi_j)<\sigma_{n-j+1}<\sigma_{n-p+1}<\delta(\psi_p)$ 的第 $p(p>j)$ 列存在，故对第 j, p 列进行同样的操作.

象这样，顺次对 $\boldsymbol{S}(s)$ 进行基本变换，经有限次之后变成列适宜矩阵 $\widetilde{\boldsymbol{S}}(s)$. 其中，$\widetilde{\boldsymbol{S}}(s)$ 的列次数是 $\sigma_n, \sigma_{n-1}, \cdots, \sigma_1$. 在各列中，主对角线元素是比其同列其它元素次数至少高 1 次的首一多项式. 对 $\widetilde{\boldsymbol{S}}(s)$ 进行第 n 列→第 1 列，第 $(n-1)$ 列→第 2 列，…，第 1 列→第 n 列的交换. 此外，令第 n 行→第 1 行，第 $(n-1)$ 列→第 2 列，…，第 1 行→第 n 行交换成的矩阵为 $\hat{\boldsymbol{S}}(s)$. $\hat{\boldsymbol{S}}(s)$ 是列适宜的，其列次数是 $\sigma_1, \sigma_2, \cdots, \sigma_n$. 而且，其主对角线元素分别是次数为 $\sigma_1, \sigma_2, \cdots, \sigma_n$ 的首一多项式. 因 $\hat{\boldsymbol{S}}(s)$ 是由 $\boldsymbol{S}(s)$ 进行基本变换得到的，则其不

变因子是 $\psi_1(s), \psi_2(s), \cdots, \psi_n(s)$. 因此，所要求的 $\boldsymbol{B}(s)$ 作为 $\boldsymbol{B}(s)=\hat{\boldsymbol{S}}(s)-\boldsymbol{A}(s)$ 给出.

证明完毕

(例) 设

$$\boldsymbol{A}(s)=\begin{bmatrix} s^2 & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 1 \end{bmatrix}, \sigma_1=2, \sigma_2=1, \sigma_3=0$$

$$\psi_1(s)=1, \psi_2(s)=1, \psi_3(s)=s^2(s+1)$$

因

$$\boldsymbol{S}(s)=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & s^2(s+1) \end{bmatrix}$$

$$\delta(\psi_1)=\sigma_3, \delta(\psi_2)<\sigma_2<\sigma_1<\delta(\psi_3)$$

由[定理 5], $\boldsymbol{S}(s)$ 变换成

$$\bar{\boldsymbol{S}}(s)=\begin{bmatrix} 1 & 0 & 0 \\ 0 & s+1 & 0 \\ 0 & 1 & s^2 \end{bmatrix}$$

因此

$$\hat{\boldsymbol{S}}(s)=\begin{bmatrix} s^2 & 1 & 0 \\ 0 & s+1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \boldsymbol{B}(s)=\begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$\boldsymbol{A}+s\boldsymbol{B}$ 的克罗内克指数[6]

在两个矩阵 $\boldsymbol{A}\in C^{n\times n}$, $\boldsymbol{B}\in C^{n\times n}$ 之间，是否能找出满足 $\boldsymbol{B}=\boldsymbol{PAP}^{-1}$ 的正则矩阵 $\boldsymbol{P}\in C^{n\times n}$ 的问题，可以通过考察 $[s\boldsymbol{I}-\boldsymbol{A}]$, $[s\boldsymbol{I}-\boldsymbol{B}]$ 形的多项式矩阵(的不变因子)得以解决. 对此加以推广，当给出 4 个矩阵 $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{A}_1$, $\boldsymbol{B}_1(\in C^{m\times n})$, 考虑是否存在同时满足

$$\boldsymbol{PAQ}=\boldsymbol{A}_1, \boldsymbol{PBQ}=\boldsymbol{B}_1 \quad (\boldsymbol{P}\in C^{m\times m}, \boldsymbol{Q}\in C^{n\times n}) \tag{33}$$

的正则矩阵 $\boldsymbol{P}$, $\boldsymbol{Q}$ 时，$\boldsymbol{A}(s)=\boldsymbol{A}+s\boldsymbol{B}$ 形的多项式矩阵具有重要意义. 多项式矩阵 $\boldsymbol{A}+s\boldsymbol{B}$ 称为阵束(pencil). 当(33)式成立，即 $\boldsymbol{P}(\boldsymbol{A}+s\boldsymbol{B})\boldsymbol{Q}=\boldsymbol{A}_1+s\boldsymbol{B}_1$ 时，称为阵束 $(\boldsymbol{A}+s\boldsymbol{B})$ 和 $(\boldsymbol{A}_1+s\boldsymbol{B}_1)$ 真正等价. 当 $m=n$ 且 $\det(\boldsymbol{A}+s\boldsymbol{B})\neq 0$ 时，称为阵束是正则的. 除此之外 ($m\neq n$, 或 $m=n$ 但 $\det(\boldsymbol{A}+s\boldsymbol{B})\equiv 0$), 称为阵束是非正则的. 在系统理论上重要的是 $m\neq n$, 即非正则的情况，以下假定 $m\neq n$.

考虑方程式

$$(\boldsymbol{A}+s\boldsymbol{B})\boldsymbol{x}(s)=\boldsymbol{0} \tag{34}$$

设 $(\boldsymbol{A}+s\boldsymbol{B})$ 的秩为 $r<n$, 则满足(34)式的多项式元素的解 $\boldsymbol{x}(s)=n\times 1(\neq\boldsymbol{0})$ 至少有一个. 在这些解中设其次数($=\boldsymbol{x}(s)$ 的元素中最高的次数)最低的为 $\boldsymbol{x}_1(s)$, 令其次数为 σ_1. 其次，在同一方程的解中，设与 $\boldsymbol{x}_1(s)$ 线性独立的解中具有最低次数 σ_2 的解为 $\boldsymbol{x}_2(s)$. 同样，在与 $\boldsymbol{x}_1(s)$, $\boldsymbol{x}_2(s)$ 线性独立的解中，设具有最低次数 σ_3 的解为 $\boldsymbol{x}_3(s)$. 因(34)式最多只有 n 个线性独立的解，则该过程经过有限次便结束，于是得到次数 $\sigma_1\leqslant\sigma_2\leqslant\sigma_3\leqslant\cdots\leqslant\sigma_p$ 的解 $\boldsymbol{x}_1(s)$, $\boldsymbol{x}_2(s)$, $\cdots$, $\boldsymbol{x}_p(s)$. $\sigma_1, \sigma_2, \cdots, \sigma_p$ 称为阵束 $(\boldsymbol{A}+s\boldsymbol{B})$ 的克罗内克列最小指数.

克罗内克行最小指数 $\eta_1, \eta_2, \cdots, \eta_q$ 可以作为 $(\boldsymbol{A}^T+s\boldsymbol{B}^T)$ 的列最小指数定义.

真正等价的两组阵束,其克罗内克指数一致.其原因是,对于(34)式的任意解 $\boldsymbol{x}(s)$,方程式 $\boldsymbol{P}(\boldsymbol{A}+s\boldsymbol{B})\boldsymbol{Q}\tilde{\boldsymbol{x}}(s)=\boldsymbol{0}$ 具有同次数的解 $\tilde{\boldsymbol{x}}(s)=\boldsymbol{Q}^{-1}\boldsymbol{x}(s)$. 反之,对于 $\boldsymbol{P}(\boldsymbol{A}+s\boldsymbol{B})\boldsymbol{Q}\tilde{\boldsymbol{x}}(s)=\boldsymbol{0}$ 的任意解 $\tilde{\boldsymbol{x}}(s)$, (34)式的同次数解 $\boldsymbol{x}(s)=\boldsymbol{Q}\tilde{\boldsymbol{x}}(s)$ 存在.

参 考 文 献

[1] F. R. Gantmacher: Theory of Matrices; Vol. 1, Chelsea(1959).

[2] 古屋:行列と行列式;培風馆(昭 32).

[3] C. C. MacDuffee: The Theory of Matrices; Chelsea (1956).

[4] W. A. Wolovich: The Application of State Feedback Invariants to Exact Model Matching; Proc. Fifth Ann. Princeton Conference, March (1971).

[5] H. H. Rosenbrock: State-Space and Multivariable Theory; John Wiley (1970).

[6] F. R. Gantmacher: Theory of Matrices; Vol. 2 Chelsea (1959).

[7] W. A. Wolovich: Linear Multivariable System; Springer-Verlag (1974).

多项式矩阵及其应用(应用部分)

在微分方程式中的应用

设 $\boldsymbol{L}(p)=m\times m$, $\boldsymbol{M}(p)=m\times r$ 分别为以微分算子 $p=\frac{d}{dt}$ 的多项式为元素的多项式矩阵,本节我们来考虑关于未知函数 $\boldsymbol{z}$ 的矩阵微分方程式

$$\boldsymbol{L}(p)\boldsymbol{z}(t)=\boldsymbol{M}(p)\boldsymbol{u}(t). \tag{1}$$

虽然可以考虑 $\boldsymbol{L}(p)$ 不是方阵的一般情况,但是描述动力学系统的方程式原则上限于方阵.

(1) 式形式的方程,是根据支配系统各部分的物理定律,考虑各部分联系时推导出来的自然形式的方程式[1]. 例如,电路中的节点电位方程及网络电流方程等就是其中的典型例子. (1)式和所谓的状态方程式比较,它直接地反映了系统的物理性质及结构. 但是另一方面,由于它不象状态方程那样经过一定数学上的处理,所以不能自动地保证解的存在性及唯一性. 而且,决定解 $\boldsymbol{z}(t)$ 必要的独立的起始条件的数目也不自明.

为了考察这一系列问题,利用前部分讲的多项式矩阵的基本变换,即可以将(1)式等价变换成便于处理形式的方程式. 在这里所谓(1)式和矩阵微分方程式

$$\widetilde{\boldsymbol{L}}(p)\boldsymbol{z}(t)=\widetilde{\boldsymbol{M}}(p)\boldsymbol{u}(t),\ \widetilde{\boldsymbol{L}}(p)=m\times m,\ \widetilde{\boldsymbol{M}}(p)=m\times r \tag{2}$$

等价,是指(1)式的任意解是(2)式的解,而且其逆亦真. 显然,当 $\widetilde{\boldsymbol{L}}(p)$ 及 $\widetilde{\boldsymbol{M}}(p)$ 分别是对 $\boldsymbol{L}(p)$, $\boldsymbol{M}(p)$ 进行基本行变换得到的矩阵时,即在

$$\widetilde{\boldsymbol{L}}(p)=\boldsymbol{T}(p)\boldsymbol{L}(p),\ \widetilde{\boldsymbol{M}}(p)=\boldsymbol{T}(p)\boldsymbol{M}(p),\ \boldsymbol{T}(p)=\text{幺模阵} \tag{3}$$

的情况下,(1)式和(2)式是等价的.

(问题 1)

对于 $\boldsymbol{L}(p)\boldsymbol{z}(t)=\boldsymbol{M}(p)\boldsymbol{u}(t)$, 考虑 $\boldsymbol{T}(p)\boldsymbol{L}(p)\boldsymbol{z}(t)=\boldsymbol{T}(p)\boldsymbol{M}(p)\boldsymbol{u}(t)$. 试证明,当 $\delta(\det\boldsymbol{T}(p))\geqslant 1$ 时,第 1 式所有的解都满足第 2 式,但是第 2 式的非自明解中有不满足第 1 式的. 而且,若 $\boldsymbol{T}(p)$ 为幺模阵,则第 1 式和第 2 式相互等价.

现在我们适当地对(1)式进行基本行变换,使之等价变换成以三角形矩阵为系数的微分方程式. 这可以按如下程序进行(它比将多项式矩阵变换成其司密斯典范形的过程简单得多). 即首先注意到第 1 列元素,如关于司密斯形所述,除了某一个元素外,使剩下的元素均变成 0, 然后将该非零元素移到(1, 1)元素. 其次,对于去掉第 1 行第 1 列后剩下的子矩阵再进行同样的操作,如此反复进行下去,即可利用基本行变换将 $\boldsymbol{L}(p)$ 变换成上三角形矩阵.

$$\left.\begin{aligned}
\tilde{l}_{11}(p)z_1(t)+\tilde{l}_{12}(p)z_2(t)+\cdots+\tilde{l}_{1m}(p)z_m(t)&=\widetilde{m}_{11}(p)u_1(t)+\cdots+\widetilde{m}_{1r}(p)u_r(t)\\
\tilde{l}_{22}(p)z_2(t)+\cdots+\tilde{l}_{2m}(p)z_m(t)&=\widetilde{m}_{21}(p)u_1(t)+\cdots+\widetilde{m}_{2r}(p)u_r(t)\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
\tilde{l}_{m-1\,m-1}(p)z_{m-1}(t)+\tilde{l}_{m-1\,m}(p)z_m(t)&=\widetilde{m}_{m-1\,1}(p)u_1(t)+\cdots+\widetilde{m}_{m-1\,r}(p)u_r(t)\\
\tilde{l}_{mm}(p)z_m(t)&=\widetilde{m}_{m1}(p)u_1(t)+\cdots+\widetilde{m}_{mr}(p)u_r(t)
\end{aligned}\right\} \tag{4}$$

1) 参看 A-I-1.

在(4)式最下一行中，当 $\tilde{l}_{mm}(p)\neq 0$ 时，可以对于 z_m 唯一解出．若 $\tilde{l}_{mm}(p)\equiv 0$，则右边不同时恒等于 0 时无解．若右边恒等于 0(即 $\tilde{m}_{m1}(p)=\tilde{m}_{m2}(p)=\cdots=\tilde{m}_{mr}(p)=0$ 或加上使 $\tilde{m}_{m1}(p)u_1(t)+\cdots\tilde{m}_{mr}(p)u_r(t)=0$ 的输入 $u_1, u_2, \cdots, u_r$)，则 z_m 可任意设定．

$z_m(t)$ 根据最下一行方程式确定后，代入其上一行．当 $\tilde{l}_{m-1\,m-1}(p)\neq 0$ 时，对于 z_{m-1} 可唯一解出．若 $\tilde{l}_{m-1\,m-1}(p)\equiv 0$，如上面关于最下一行所述，若其它项之和不恒等于 0 则无解，若等于 0，则 z_{m-1} 可任意设定．

如此进行下去，直到第$(m-2)$，第$(m-3)$，…，第 1 行为止，最后可以得到，(4)式具有唯一解 $z_1(t), z_2(t), \cdots, z_m(t)$ 的充分和必要条件是，$\tilde{l}_{11}(p), \tilde{l}_{22}(p), \cdots, \tilde{l}_{mm}(p)$ 均为不恒等于零的多项式．因(4)式是由(1)式进行基本行变换后得到的，因而它和 $\det\boldsymbol{L}(p)\neq 0$ 的条件是一样的．归纳以上得如下定理．

[定理 1]

“(1)式(对于任意输入)具有唯一解的充分和必要条件是 $\det\boldsymbol{L}(p)\neq 0$.”

在实现问题中的应用

所谓实现问题，是根据与系统输入输出响应有关的信息，决定系统的状态方程式 $\dot{\boldsymbol{x}}(t)=\boldsymbol{A}\boldsymbol{x}(t)+\boldsymbol{B}\boldsymbol{u}(t)$ 和输出方程式 $\boldsymbol{y}(t)=\boldsymbol{C}\boldsymbol{x}(t)+\boldsymbol{D}\boldsymbol{u}(t)$．根据输入输出信息用什么形式给出，可以考虑各种各样的实现问题，但是这里作为前节的续，以微分方程式 $\boldsymbol{L}(p)\boldsymbol{z}(t)=\boldsymbol{M}(p)\boldsymbol{u}(t)$ 为对象，首先讨论一下下列关于传递函数的实现问题．

(i) $\boldsymbol{L}(p)\boldsymbol{z}(t)=\boldsymbol{M}(p)\boldsymbol{u}(t)$ 的实现

我们假定，在系统的输入 $\boldsymbol{u}$ 和适当选择的物理变量 $\boldsymbol{z}=(z_1, z_2, \cdots, z_m)^T$ 之间，(1)式微分方程式成立，当然认为 $\det\boldsymbol{L}(p)\neq 0$．并且，假定 $\boldsymbol{z}$ 本身就是系统的输出 $\boldsymbol{y}$，即 $\boldsymbol{y}=\boldsymbol{z}$．在 $\boldsymbol{y}$ 用 $\boldsymbol{z}$ 和 $\boldsymbol{u}$ 的线性组合 $\boldsymbol{y}(t)=\boldsymbol{R}\boldsymbol{z}(t)+\boldsymbol{S}\boldsymbol{u}(t)$ 表示的情况下也可以同样地处理．

为了能使用后面的算法，首先假定，对(1)式进行基本行变换 $\boldsymbol{T}(p)$，而所得到新的系数矩阵 $\tilde{\boldsymbol{L}}(p)=\boldsymbol{T}(p)\boldsymbol{L}(p)$ 满足下列条件．

(i) $\tilde{\boldsymbol{L}}(p)=[\tilde{l}_{ij}(p)]$ 是行适宜的，而且

$$\delta(\tilde{l}_{ii}(p))>\delta(\tilde{l}_{ij}(p)) \quad j\neq i,\ i=1, 2, \cdots, m \tag{5}$$

为了使输入输出关系 $\boldsymbol{L}(p)\boldsymbol{z}(t)=\boldsymbol{M}(p)\boldsymbol{u}(t)$，$\boldsymbol{y}(t)=\boldsymbol{z}(t)$ 有等价的标准形状态方程式 $\dot{\boldsymbol{x}}(t)=\boldsymbol{A}\boldsymbol{x}(t)+\boldsymbol{B}\boldsymbol{u}(t)$ 和输出方程式 $\boldsymbol{y}(t)=\boldsymbol{C}\boldsymbol{x}(t)+\boldsymbol{D}\boldsymbol{u}(t)$，至少必须输入输出之间的传递函数是适宜的．因此第二个假定

(ii) 设

$$\boldsymbol{L}^{-1}(s)\boldsymbol{M}(s)\,(=\tilde{\boldsymbol{L}}^{-1}(p)\tilde{\boldsymbol{M}}(p))\to\boldsymbol{D}\quad(=\text{常数矩阵}),\ s\to\infty \tag{6}$$

成立，式中 $\tilde{\boldsymbol{M}}(p)=\boldsymbol{T}(p)\boldsymbol{M}(p)$．

根据条件(i)，(ii)和前部分讲过的引理 5，注意对于 $\tilde{\boldsymbol{M}}(p)$ 的各行

$$\delta(\tilde{m}_{ij}(p))\leqslant\delta(\tilde{l}_{ii}(p)) \tag{7}$$

成立．

现在先给出对于下面要讲的实现算法所必需的引理．

[引理 1]

“n 阶纯量微分方程式

$$(a_np^n+a_{n-1}p^{n-1}+\cdots+a_0)z(t)=(b_np^n+b_{n-1}p^{n-1}+\cdots+b_0)u(t),\quad a_n\neq 0$$

的最小维数实现[1)](之一)可给出如下

$$\left.\begin{aligned}\dot{\boldsymbol{x}}(t)&=\frac{1}{a_n}\begin{bmatrix}-a_{n-1} & a_n & 0 & \cdots & 0\\ -a_{n-2} & 0 & a_n & \cdots & 0\\ \vdots & \vdots & & & \\ -a_1 & 0 & 0 & & a_n\\ -a_0 & 0 & 0 & & 0\end{bmatrix}\boldsymbol{x}(t)+\frac{1}{a_n}\begin{bmatrix}a_nb_{n-1} & -a_{n-1}b_n\\ a_nb_{n-2} & -a_{n-2}b_n\\ \cdots\cdots & \cdots\cdots\\ a_nb_0 & -a_0b_n\end{bmatrix}u(t)\\ z(t)&=\left[\frac{1}{a_n}\,0\cdots 0\right]\boldsymbol{x}(t)+\frac{b_n}{a_n}u(t)\end{aligned}\right\}$$

其中状态变量选择如下

$$x_1(t)=a_nz(t)-b_nu(t)$$
$$\begin{aligned}x_2(t)&=a_nz^{(1)}(t)-b_nu^{(1)}(t)+a_{n-1}z(t)-b_{n-1}u(t)\\ &=\dot{x}_1(t)+a_{n-1}z(t)-b_{n-1}u(t)\end{aligned}$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\begin{aligned}x_n(t)&=a_nz^{(n-1)}(t)-b_nu^{(n-1)}(t)+a_{n-1}z^{(n-2)}(t)-b_{n-1}u^{(n-2)}(t)+\cdots+a_1z(t)-b_1u(t)\\ &=\dot{x}_{n-1}(t)-a_1z(t)-b_1u(t)\text{"}\end{aligned}$$

[引理 2]

"两输入 n 阶纯量微分方程式

$$\begin{aligned}&(a_np^n+a_{n-1}p^{n-1}+\cdots+a_0)z(t)\\ &\quad=(b_np^n+b_{n-1}p^{n-1}+\cdots+b_0)u_1(t)+(c_np^n+c_{n-1}p^{n-1}+\cdots+c_0)u_2(t)\end{aligned}$$

的最小实现亦可以参照引理 1 的结果,即若选取状态变量

$$\begin{aligned}x_1(t)&=a_nz(t)-b_1u_1(t)-c_nu_2(t)\\ x_2(t)&=\dot{x}_1(t)+a_{n-1}z(t)-b_{n-1}u_1(t)-c_{n-1}u_2(t)\\ &\cdots\cdots\cdots\cdots\cdots\cdots\\ x_n(t)&=\dot{x}_{n-1}(t)+a_1z(t)-b_1u_1(t)-c_1u_2(t),\end{aligned}$$

则状态方程式和输出方程式的矩阵 $\boldsymbol{A}$ 及 $\boldsymbol{C}$ 和(7)式相同,而 $\boldsymbol{B}$ 及 $\boldsymbol{D}$ 部分分别为

$$\boldsymbol{B}=\frac{1}{a_n}\begin{bmatrix}a_nb_{n-1}-a_{n-1}b_n & a_nc_{n-1}-a_{n-1}c_n\\ a_nb_{n-2}-a_{n-2}b_n & a_nc_{n-2}-a_{n-2}c_n\\ \cdots\cdots & \cdots\cdots\\ a_nb_0-a_0b_n & a_nc_0-a_0c_n\end{bmatrix},\quad \boldsymbol{D}=\frac{1}{a_n}[b_nc_n]\text{"}$$

因引理 1 是大家很熟悉的结果[1,2],而且引理 2 是引理 1 的直接推广,故这里证明省略.

基于多项式矩阵理论,(1)式微分方程式的最小实现算法可叙述如下.

"(第 1 步)

对 $\boldsymbol{L}(p)\boldsymbol{z}(t)=\boldsymbol{M}(p)\boldsymbol{u}(t)$, $\boldsymbol{y}(t)=\boldsymbol{z}(t)$ 进行基本行变换,认为变换后的式子 $\widetilde{\boldsymbol{L}}(p)\boldsymbol{z}(t)=\widetilde{\boldsymbol{M}}(p)\boldsymbol{u}(t)$ 满足上述条件(i),(ii).

(第 2 步)

在 $\widetilde{\boldsymbol{L}}(p)\boldsymbol{z}(t)=\widetilde{\boldsymbol{M}}(p)\boldsymbol{u}(t)$ 的第 1 行

1) 因是可观测的,故为最小实现.

$$\tilde{l}_{11}(p)z_1(t)=-\tilde{l}_{12}(p)z_2(t)-\cdots$$
$$-\tilde{l}_{1m}(p)z_m(t)+\widetilde{m}_{11}(p)u_1(t)+\cdots+\widetilde{m}_{1r}(p)u_r(t)$$

中，将 $z_2, z_3, \cdots, z_m, u_1, \cdots, u_r$ 看成输入，利用引理 2[1]求最小实现．这时，必需 $d_1 \triangleq \delta(\tilde{l}_{11}(p))$ 个状态变量 $\boldsymbol{x}_1=(x_{11}, x_{21}, \cdots, x_{d1})^T$.

同样，对于第 2 行也将 $z_1, z_3, \cdots, z_m, u_1, \cdots, u_r$ 看作输入，引入 $d_2 \triangleq \delta(\tilde{l}_{22}(p))$ 个状态变量求最小实现．以下，对第 3 行，…，第 m 行亦进行同样操作．

根据以上结果，可以得到包含有 $d \triangleq d_1+d_2+\cdots+d_m=\delta(\det\tilde{\boldsymbol{L}}(p))=\delta(\det \boldsymbol{L}(p))$ 个状态变量的 m 个状态方程式及输出方程式．特别要注意的是，为使(5)式成立，输出方程式变成

$$z_i(t)=\boldsymbol{C}_i\boldsymbol{x}_i(t)+\boldsymbol{D}_i\boldsymbol{u}(t),\ i=1,\ 2,\ \cdots,\ m \tag{8}$$

$z_j(t)\ (j\neq i)$ 不以显函数的形式出现．

（第 3 步）

在第 2 步求得的各状态方程中，将输出方程式 (8) 分别代入包含在输入项中的 $z_1, z_2, \cdots, z_m$，消去 $z_1, z_2, \cdots, z_m$，整理后即可得到标准形状态方程式．”

（例 1）

第 1 步

$$\begin{bmatrix} p^2+3p+1 & 2p+3 \\ p^3+3p^2+p & 3p^2+3p+6 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}=\begin{bmatrix} 1 & 0 \\ p+1 & p+3 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}\longrightarrow$$

将第 1 行 ×(−p) 加到第 2 行

$$\begin{bmatrix} p^2+3p+1 & 2p+3 \\ 0 & p^2+6 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}=\begin{bmatrix} 1 & 0 \\ 1 & p+3 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

第 2 步

第 1 行：$(p^2+3p+1)z_1=-(2p+3)z_2+u_1$

$$\left.\begin{aligned} x_{11}&=z_1 \\ x_{21}&=\dot{z}_1+3z_1+2z_2 \end{aligned}\right\}\Rightarrow\begin{bmatrix} \dot{x}_{11} \\ \dot{x}_{21} \end{bmatrix}$$
$$=\begin{bmatrix} -3 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} x_{11} \\ x_{21} \end{bmatrix}+\begin{bmatrix} -2 \\ -3 \end{bmatrix}z_2+\begin{bmatrix} 0 \\ 1 \end{bmatrix}u_1$$

第 2 行：$(p^2+6)z_2=u_1+(p+3)u_2$

$$\left.\begin{aligned} x_{12}&=z_2 \\ x_{22}&=\dot{z}_2-u_2 \end{aligned}\right\}\Rightarrow\begin{bmatrix} \dot{x}_{12} \\ \dot{x}_{22} \end{bmatrix}$$
$$=\begin{bmatrix} 0 & 1 \\ -6 & 0 \end{bmatrix}\begin{bmatrix} x_{12} \\ x_{22} \end{bmatrix}+\begin{bmatrix} 0 \\ 1 \end{bmatrix}u_1+\begin{bmatrix} 1 \\ 3 \end{bmatrix}u_2$$

第 3 步

整理后得

$$\begin{bmatrix} \dot{x}_{11} \\ \dot{x}_{21} \\ \dot{x}_{12} \\ \dot{x}_{22} \end{bmatrix}=\begin{bmatrix} -3 & 1 & -2 & 0 \\ -1 & 0 & -3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -6 & 0 \end{bmatrix}\begin{bmatrix} x_{11} \\ x_{21} \\ x_{12} \\ x_{22} \end{bmatrix}+\begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

1) 对于 3 个以上输入的实现，按引理 2 很容易求出．

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} x_{11} \\ x_{12} \end{bmatrix}$$

(ii) 传递函数的实现

设给出以既约实系数有理函数为元素的传递函数(矩阵) $\boldsymbol{H}(s)=[h_{ij}(s)]=m\times r$, 并假定 $\boldsymbol{H}(s)$ 是真正适宜的, 即 $\boldsymbol{H}(\infty)=\boldsymbol{0}$. 在 $\boldsymbol{H}(s)$ 仅仅适宜 ($\boldsymbol{H}(\infty)=\boldsymbol{D}\neq\boldsymbol{0}$) 的情况下, 以 $\hat{\boldsymbol{H}}(s)=\boldsymbol{H}(s)-\boldsymbol{H}(\infty)$ 为对象即可. $\boldsymbol{H}(\infty)$ 对 $\hat{\boldsymbol{H}}(s)$ 的最小实现 $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ 仅附带地影响 $\boldsymbol{D}(=\boldsymbol{H}(\boldsymbol{0}))$ 矩阵.

我们知道, 求这种 $\boldsymbol{H}(s)$ 的最小实现的方法很多[3]. 在这里, 为了弄清系统动力学信息以何种形式包含在传递函数之中, 我们主要讨论很方便的卡尔曼-海曼(Kalman-Heymann)实现法[4]. 这就是根据下列步骤求 $\boldsymbol{H}(s)$ 的最小实现的方法.

“(第 1 步)

首先, 将 $\boldsymbol{H}(s)$ 变换成司密斯-麦克米兰典范形. 即令 $\alpha_1(s)$ 为 $\boldsymbol{H}(s)$ 各元素分母最小公倍首一多项式, 因 $\alpha_1(s)\boldsymbol{H}(s)$ 为多项式矩阵, 则 $\alpha_1(s)\boldsymbol{H}(s)$ 变换成下列司密斯典范形

$$\alpha_1(s)\boldsymbol{H}(s)=\boldsymbol{P}(s)\begin{bmatrix} \gamma_1(s) & & & \\ & \gamma_2(s) & \ddots & \boldsymbol{0} \\ \boldsymbol{0} & & \gamma_R(s) & \boldsymbol{0} \end{bmatrix}\boldsymbol{Q}(s) \tag{9}$$

式中 $R=\operatorname{rank}\alpha_1(s)\boldsymbol{H}(s), \gamma_1(s), \cdots, \gamma_R(s)$ 是 $\alpha_1(s)\boldsymbol{H}(s)$ 的不变因子, $\boldsymbol{P}(s)=m\times m, \boldsymbol{Q}(s)=r\times r$ 均为幺模多项式矩阵.

将(9)式两边除以 $\alpha_1(s)$, 若 $\gamma_i(s)/\alpha_1(s)$ 的分母分子有公因子, 将其消去, 并写成 $\gamma_i(s)/\alpha_1(s)=\tilde{n}_i(s)/d_i(s)$ 的既约形式.

这时由(9)式传递函数 $\boldsymbol{H}(s)$ 可表示成

$$\boldsymbol{H}(s)=\boldsymbol{P}(s)\begin{bmatrix} \dfrac{\tilde{n}_1(s)}{d_1(s)} & & \\ & \ddots & \\ & & \dfrac{\tilde{n}_R(s)}{d_R(s)} & \boldsymbol{0} \\ \boldsymbol{0} & & & \boldsymbol{0} \end{bmatrix}\boldsymbol{Q}(s). \tag{10}$$

(10)式称为 $\boldsymbol{H}(s)$ 的司密斯-麦克米兰典范形.

此外, 多项式

$$d(s)\triangleq d_1(s),\ d_2(s),\ \cdots,\ d_R(s) \tag{11}$$

的次数 $\delta_M=\delta(d(s))$ 称为 $\boldsymbol{H}(s)$ 的麦克米兰次数.

(第 2 步)

在 $\boldsymbol{H}(s)$ 的司密斯-麦克米兰典范形(10)式中, 取出对角项 $\tilde{n}_i(s)/d_i(s)$ 的真正适宜部分. 即当 $\delta(\tilde{n}_i(s))\geqslant\delta(d_i(s))$ 时, 用 $d_i(s)$ 除 $\tilde{n}_i(s)$, 得

$$\tilde{n}_i(s)=q_i(s)d_i(s)+n_i(s),\ i=1,\ 2,\ \cdots,\ R$$

其中设余数为 $n_i(s)$, 则 $n_i(s)/d_i(s)$ 是 $\tilde{n}_i(s)/d_i(s)$ 的真正适宜部分1). 若 $\delta(\tilde{n}_i(s))<\delta(d_i(s))$, 则 $\tilde{n}_i(s)/d_i(s)$ 保持原样即可. 将

$$g_i(s)\triangleq\frac{n_i(s)}{d_i(s)}\triangleq\frac{b_{n_i-1i}s^{n_i-1}+\cdots+b_{0i}}{s^{n_i}+a_{n_i-1}s^{n_i-1}+\cdots+a_{0i}} \tag{12}$$

1) 即真分式部分. ——译者注

为纯量传递函数，求出这个最小实现，设为$(\tilde{\boldsymbol{A}}_i, \tilde{\boldsymbol{b}}_i, \tilde{\boldsymbol{c}}_i)$. 例如，作为最小实现，可控标准形

$$\tilde{\boldsymbol{A}}_i=\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_{0i} & -a_{1i} & & \cdots & -a_{n_i-1i} \end{bmatrix}, \quad \tilde{\boldsymbol{b}}_i=\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix},$$

$$\tilde{\boldsymbol{c}}_i=[b_{0i}, b_{1i}, \cdots, b_{n_i-1i}] \tag{13}$$

是很方便的.

(第3步)

利用在第2步中求出的R个最小实现$(\tilde{\boldsymbol{A}}_i, \tilde{\boldsymbol{b}}_i, \tilde{\boldsymbol{c}}_i)$, $i=1, 2, \cdots, R$, 定义

$$\tilde{\boldsymbol{A}}\triangleq\begin{bmatrix} \tilde{\boldsymbol{A}}_1 & & & \\ & \tilde{\boldsymbol{A}}_2 & & \\ & & \ddots & \\ & & & \tilde{\boldsymbol{A}}_R \end{bmatrix}\in R^{\delta_M\times\delta_M} \tag{14}$$

$$\tilde{\boldsymbol{B}}\triangleq\begin{bmatrix} \tilde{\boldsymbol{b}}_1 & & & \\ & \tilde{\boldsymbol{b}}_2 & & \\ & & \ddots & \boldsymbol{0} \\ & & & \tilde{\boldsymbol{b}}_R \end{bmatrix}\in R^{\delta_M\times r}, \quad \tilde{\boldsymbol{C}}\triangleq\begin{bmatrix} \tilde{\boldsymbol{c}}_1 & & & \\ & \tilde{\boldsymbol{c}}_2 & & \\ & & \ddots & \\ & & \boldsymbol{0} & \tilde{\boldsymbol{c}}_R \end{bmatrix}\in R^{m\times\delta_M} \tag{15}$$

将(9)式中的$\boldsymbol{P}(s)$, $\boldsymbol{Q}(s)$展开成s的幂级数形式，设其分别为

$$\boldsymbol{P}(s)=\sum_{i=1}^{l_1}\boldsymbol{P}_i s^i, \quad \boldsymbol{Q}(s)=\sum_{i=1}^{l_2}\boldsymbol{Q}_i s^i, \quad (\boldsymbol{P}_i, \boldsymbol{Q}_i=\text{常数矩阵})$$

则

$$\left(\boldsymbol{A}\triangleq\tilde{\boldsymbol{A}}, \ \boldsymbol{B}\triangleq\sum_{i=1}^{l_2}\tilde{\boldsymbol{A}}^i\tilde{\boldsymbol{B}}\boldsymbol{Q}_i, \ \boldsymbol{C}\triangleq\sum_{i=1}^{l_1}\boldsymbol{P}_i\tilde{\boldsymbol{C}}\tilde{\boldsymbol{A}}^i\right) \tag{16}$$

是$\boldsymbol{H}(s)$的最小实现.”

以后再来检验该结果是否妥当. 注意，作为其直接结果，对于$\boldsymbol{H}(s)$和其最小实现的关系，可以得出若干有趣的事实.

(注意)

(a) 由(11), (14), (16)式可见

最小实现的维数$=\boldsymbol{H}(s)$的麦克米兰次数

(b) 如前部分(例3)所示，$[s\boldsymbol{I}-\tilde{\boldsymbol{A}}_i]$的不变因子(除去自明项$\gamma(s)=1$)是$d_i(s)=s^{n_i}+a_{n_i-1i}s^{n_i-1}+\cdots+a_{0i}$. 在求$\boldsymbol{H}(s)$的司密斯-麦克米兰典范形(10)式的过程中，因$\gamma_i(s)$, $i=1, 2, \cdots, R$是$a_1(s)\boldsymbol{H}(s)$的不变因子，则$\gamma_{i+1}(s)$可被$\gamma_i(s)$整除，即

$$\gamma_{i+1}(s)/\gamma_i(s)=(\tilde{n}_{i+1}(s)/\tilde{n}_i(s))(d_i(s)/d_{i+1}(s))$$

是多项式，而且这表明$d_i(s)$可被$d_{i+1}(s)$整除. 以上及$\boldsymbol{A}(=\tilde{\boldsymbol{A}})$是分块对角形，意味着$[s\boldsymbol{I}-\boldsymbol{A}]$的不变因子(再除去自明项)是$d_R(s)$, $d_{R-1}(s)$, $\cdots$, $d_1(s)$. 因为$\boldsymbol{H}(s)$的任意最小实现的矩阵$\boldsymbol{A}$和用上述卡尔曼-海曼方法求得的$\boldsymbol{A}$((14)式)相似[1)]，则一般得到下列结论.

---

1) 参看A-I-11中系7.

对于最小实现的矩阵 $\boldsymbol{A}$,

$$[s\boldsymbol{I}-\boldsymbol{A}]\text{的不变因子}=\{d_R(s),\ d_{R-1}(s),\ \cdots,\ d_1(s)\} \tag{17}$$

特别是

$$[s\boldsymbol{I}-\boldsymbol{A}]\text{的特征多项式}=d_1(s)\quad d_2(s)\cdots d_R(s) \tag{18}$$

$$[s\boldsymbol{I}-\boldsymbol{A}]\text{的最小多项式}=d_1(s) \tag{19}$$

(c) 在推导司密斯-麦克米兰典范形时,假定成 $\gamma_1(s)$ 和 $\alpha_1(s)$ 之间有公因子的情况,根据 $\gamma_1(s)/\alpha_1(s)=\tilde{n}_1(s)/d_1(s)$ 定出 $d_1(s)$,实际上公因子不存在. 因此,$d_1(s)=\alpha_1(s)$. 这是因为,若在(10)式两边乘以 $d_1(s)$,则 $d_1(s)\boldsymbol{H}(s)$ 变成多项式矩阵,若 $\delta(d_1(s))<\delta(\alpha_1(s))$,则与 $\alpha_1(s)$ 是 $\boldsymbol{H}(s)$ 各元素最小公倍分母的假定相矛盾. 即

$$\alpha_1(s)=d_1(s)=\boldsymbol{H}(s)\text{的全元素最小公倍分母多项式} \tag{20}$$

陈(Chen)[6]证明,将(20)式结果推广,令

$$\alpha_k(s)\triangleq\boldsymbol{H}(s)\text{的所有 }k\text{ 级以及 }k\text{ 级以下子式的最小公倍分母(首一)多项式},$$
$$k=1,\ 2,\ \cdots,\ \min(m,\ r)$$

时,则

$$\alpha_k(s)=d_1(s)d_2(s)\cdots d_k(s)\quad k=1,\ 2,\ \cdots,\ \min(m,\ r) \tag{21}$$

例如,利用该公式,最小实现的特征多项式作为 $\alpha_{\min(m,r)}(s)$,可直接由 $\boldsymbol{H}(s)$ 求出.

(例 2)

$$\boldsymbol{H}(s)=\begin{bmatrix}\dfrac{1}{s} & \dfrac{s+1}{s^2}\\[2ex] \dfrac{1}{s^2} & \dfrac{1}{s+1}\end{bmatrix},\quad \begin{array}{l}\alpha_1(s)=s^2(s+1),\ \alpha_2(s)=s^4(s+1),\\ d_1(s)=s^2(s+1),\ d_2(s)=s^2\end{array}$$

(卡尔曼-海曼算法的证明)

以上讨论的卡尔曼-海曼算法可以作为下列引理的直接应用证明.

[引理 3]

“由幺模多项式矩阵 $\boldsymbol{P}(s)$, $\boldsymbol{Q}(s)$ 的系数矩阵按照(16)式关系联系的两个系统 $S=(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{C})$ 和 $\tilde{S}=(\tilde{\boldsymbol{A}},\ \tilde{\boldsymbol{B}},\ \tilde{\boldsymbol{C}})$ 之间,下列性质成立.

(i) $d_1(s)\boldsymbol{C}[s\boldsymbol{I}-\boldsymbol{A}]^{-1}\boldsymbol{B}=\boldsymbol{P}(s)\ (d_1(s)\tilde{\boldsymbol{C}}[s\boldsymbol{I}-\tilde{\boldsymbol{A}}]^{-1}\tilde{\boldsymbol{B}})\boldsymbol{Q}(s),\ (\bmod d_1(s))$[1)]

(ii) S=可控,可观测$\Leftrightarrow\tilde{S}$=可控且,

其可观测式中 $d_1(s)$ 是 $\boldsymbol{A}=\tilde{\boldsymbol{A}}$ 的最小多项式.”

根据上述注意(c),因 $d_1(s)$ 是 $\boldsymbol{A}=\tilde{\boldsymbol{A}}$ 的最小多项式,则可用引理 3(i),利用它对于(16)式系统$(\boldsymbol{A},\ \boldsymbol{B},\ \boldsymbol{C})$,得

$$d_1(s)\boldsymbol{C}[s\boldsymbol{I}-\boldsymbol{A}]^{-1}\boldsymbol{B}=d_1(s)\boldsymbol{P}(s)\tilde{\boldsymbol{C}}[s\boldsymbol{I}-\tilde{\boldsymbol{A}}]^{-1}\tilde{\boldsymbol{B}}\boldsymbol{Q}(s),\ \bmod d_1(s)$$

$$=d_1(s)\boldsymbol{P}(s)\begin{bmatrix}\dfrac{n_1(s)}{d_1(s)} & & & \\ & \ddots & & \\ & & \dfrac{n_R(s)}{d_R(s)} & \boldsymbol{0}\\ & \boldsymbol{0} & & \boldsymbol{0}\end{bmatrix}\boldsymbol{Q}(s)$$

1) 两个多项式矩阵 $\boldsymbol{A}(s)$ 和 $\boldsymbol{B}(s)$ 按 $\bmod q(s)$ 相等是指,其差能用 $q(s)$ 整除,即有满足 $\boldsymbol{A}(s)=\boldsymbol{B}(s)+\boldsymbol{C}(s)q(s)$ 的多项式矩阵 $\boldsymbol{C}(s)$ 存在.

$$=d_1(s)\boldsymbol{P}(s)\begin{bmatrix}\frac{n_1(s)}{d_1(s)} & & & \\ & \ddots & & \\ & & \frac{n_R(s)}{d_R(s)} & \mathbf{0} \\ \mathbf{0} & & & \mathbf{0}\end{bmatrix}\boldsymbol{Q}(s)$$

$$+d_1(s)\boldsymbol{P}(s)\begin{bmatrix}q_1(s) & & & \\ & \ddots & & \\ & & q_R(s) & \mathbf{0} \\ & \mathbf{0} & & \mathbf{0}\end{bmatrix}\boldsymbol{Q}(s)$$

$$(\bmod\, d_1(s))$$

$$=d_1(s)\boldsymbol{P}(s)\begin{bmatrix}\frac{\tilde{n}_1(s)}{d_1(s)} & & & \\ & \ddots & & \\ & & \frac{\tilde{n}_R(s)}{d_R(s)} & \mathbf{0} \\ \mathbf{0} & & & \mathbf{0}\end{bmatrix}\boldsymbol{Q}(s)=d_1(s)\boldsymbol{H}(s) \tag{22}$$

但是，因 $\boldsymbol{C}[s\boldsymbol{I}-\boldsymbol{A}]^{-1}\boldsymbol{B}$ 及 $\boldsymbol{H}(s)$ 均为真正适宜的有理函数矩阵，则(22)式成立意味着 $\boldsymbol{C}[s\boldsymbol{I}-\boldsymbol{A}]^{-1}\boldsymbol{B}=\boldsymbol{H}(s)$．由此可见，(16)式是 $\boldsymbol{H}(s)$ 的实现.

此外，因 $(\tilde{\boldsymbol{A}}_i, \tilde{\boldsymbol{b}}_i, \tilde{\boldsymbol{c}}_i)$ 是 $g_i(s)$ 的最小实现，则可控且可观测．根据(14)，(15)式，系统 $(\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{B}}, \tilde{\boldsymbol{C}})$ 是子系统 $(\tilde{\boldsymbol{A}}_i, \tilde{\boldsymbol{b}}_i, \tilde{\boldsymbol{c}}_i)$，$i=1, 2, \cdots, R$ 的直和，因而也和各子系统一样，可控且可观测[1]．根据引理 3(ii)，这意味着 $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ 可控且可观测．因此，证明 $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ 是 $\boldsymbol{H}(s)$ 的最小实现.

在传递函数分解形中的应用

(i) 分解形及其性质

将前节(ii)的结果用于适宜的纯量传递函数 $\hat{h}(s)=N(s)/D(s)$，可以看到，当分子 $N(s)$ 和分母 $D(s)$ 是相互既约的多项式时，其最小实现的维数与 $D(s)$ 的维数一致，$D(s)$ 给出特征多项式的常数倍．对于矩阵传递函数，我们来研究一下与此相当的性质．为此，我们考虑将 $\boldsymbol{H}(s)=m\times r$ 分解成

$$\boldsymbol{H}(s)=\boldsymbol{N}(s)\boldsymbol{D}(s)^{-1}$$

$$\boldsymbol{N}(s)=m\times r,\ \boldsymbol{D}(s)=r\times r \text{ 是多项式矩阵} \tag{23}$$

或

$$\boldsymbol{H}(s)=\boldsymbol{D}(s)^{-1}\boldsymbol{N}(s)$$

$$\boldsymbol{N}(s)=m\times r,\ \boldsymbol{D}(s)=m\times m \text{ 是多项式矩阵.} \tag{24}$$

我们将(23)式称为 $\boldsymbol{H}(s)$ 的右分解形，(24)式称为 $\boldsymbol{H}(s)$ 的左分解形．特别是，当(23)[(24)]式中 $\boldsymbol{D}(s)$ 和 $\boldsymbol{N}(s)$ 是相互右[左]素矩阵[2]时，称为右[左]既约分解.

对于给定的真正适宜的 $\boldsymbol{H}(s)$，其右既约分解（之一）可以从司密斯-麦克米兰典范形直接得到．即若令

1) 参照 A-I-11 中(系 6).

2) 参看前部分定义 3.

$$N_R(s) \triangleq P(s)\begin{bmatrix} n_1(s) & & & & \\ & \ddots & & \mathbf{0} & \\ & & n_R(s) & & \\ & & & 0 & \\ & \mathbf{0} & & & \ddots \\ & & & & & 0 \end{bmatrix}, \quad D_R(s) \triangleq Q^{-1}(s)\begin{bmatrix} d_1(s) & & & & \\ & \ddots & & \mathbf{0} & \\ & & d_R(s) & & \\ & & & 1 & \\ & \mathbf{0} & & & \ddots \\ & & & & & 1 \end{bmatrix} \tag{25}$$

显然，$H(s)=N_R(s)D_R(s)^{-1}$，式中$N_R(s)$和$D_R(s)$是相互右素的．其原因是，$d_i(s)$和$n_i(s)$是既约的，对于某个多项式$p_i(s)$，$q_i(s)$，$q_i(s)d_i(s)+p_i(s)n_i(s)=1(i=1, 2, \cdots, R)$成立．因此，对于多项式矩阵

$$\tilde{Q}(s)=\begin{bmatrix} q_1(s) & & & & \\ & \ddots & & & \\ & & q_R(s) & & \\ & & & 1 & \\ & & & & \ddots \\ & & & & & 1 \end{bmatrix}Q(s)$$

$$\tilde{P}(s)=\begin{bmatrix} p_1(s) & & & & \\ & \ddots & & & \\ & & p_R(s) & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix}P(s)^{-1}$$

下式成立[1)]

$$\tilde{Q}(s)D_R(s)+\tilde{P}(s)N_R(s)=I$$

右既约分解也可以不用司密斯-麦克米兰典范形，而用更一般的方法求出．对于给定的$H(s)$（可以不一定是真正适宜的），设其全元素最小公倍分母为$\alpha_1(s)$，则可写成

$$H(s)=N(s)\frac{1}{\alpha_1(s)}, \quad N(s)\text{是多项式矩阵} \tag{26}$$

若令$D(s)\triangleq\alpha_1(s)I$，则$N(s)D(s)^{-1}$是$H(s)$的右分解．求出$N(s)$和$D(s)$的(一个)gcrd $G(s)$（例如，根据麦克达菲算法），令

$$N(s)=\tilde{N}(s)G(s), \quad D(s)=\tilde{D}(s)G(s) \tag{27}$$

则$\tilde{N}(s)\tilde{D}(s)^{-1}$是$H(s)$的右既约分解(之一)．

利用该算法，可以比用司密斯-麦克米兰典范形较少的计算量得到右既约分解．

这样给出的传递函数$H(s)$的右分解不是唯一的．那么在$H(s)$的两个不同的分解之间是否有什么关系呢？对此，波波夫(Popov)证明了下列重要结果．

［定理 2］ 设$H(s)=m\times r$是以实系数有理函数为元素的传递函数，它的一个右[左]既约分解为$H(s)=N_1(s)D_1(s)^{-1}[=D_2(s)^{-1}N_2(s)]$．

(a) 对于某个正则多项式矩阵$R_1(s)=r\times r[R_2(s)=m\times m]$，

$\tilde{N}_1(s)=N_1(s)R_1(s)$，$\tilde{D}_1(s)=D_1(s)R_1(s)$ $[\tilde{N}_2(s)=R_2(s)N_2(s)$，$\tilde{D}_2(s)=R_2(s)D_2(s)]$成立$\Leftrightarrow\tilde{N}_1(s)\tilde{D}_1(s)^{-1}[\tilde{D}_2(s)^{-1}\tilde{N}_2(s)]$是$H(s)$的一个右[左]分解

(b) 对于某个幺模多项式矩阵$R_1(s)[R_2(s)]$，

$\tilde{N}_1(s)=N_1(s)R_1(s)$，$\tilde{D}_1(s)=D_1(s)R_1(s)$ $[\tilde{N}_2(s)=R_2(s)N_2(s)$，$\tilde{D}_2(s)=R_2(s)D_2(s)]$成立$\Leftrightarrow\tilde{N}(s)\tilde{D}(s)^{-1}[\tilde{D}_2{}^{-1}(s)\tilde{N}_2(s)]$是$H(s)$的一个右[左]既约分解

1) 参看前部分定理 4．

（证明） 这里只证明右分解，而对于左分解也是同样的．为了书写简单，略去注角1．(a)⇒显然．现来证明⇐：

因 $\boldsymbol{N}(s)\boldsymbol{D}(s)^{-1}=\tilde{\boldsymbol{N}}(s)\tilde{\boldsymbol{D}}(s)^{-1}$，则$\tilde{\boldsymbol{N}}(s)=\boldsymbol{N}(s)\{\boldsymbol{D}(s)^{-1}\tilde{\boldsymbol{D}}(s)\}$．而且，$\tilde{\boldsymbol{D}}(s)=\boldsymbol{D}(s)\{\boldsymbol{D}(s)^{-1}\tilde{\boldsymbol{D}}(s)\}$自明．因此，若能证明 $\boldsymbol{D}(s)^{-1}\tilde{\boldsymbol{D}}(s)$ 是多项式矩阵则是充分的．$\boldsymbol{D}(s)$ 和 $\boldsymbol{N}(s)$ 是相互右素的，即

$$\begin{bmatrix}\boldsymbol{D}(s)\\ \boldsymbol{N}(s)\end{bmatrix}=\boldsymbol{P}(s)\begin{bmatrix}\boldsymbol{I}_r\\ \boldsymbol{0}\end{bmatrix}\boldsymbol{Q}(s),\ \boldsymbol{P}(s),\ \boldsymbol{Q}(s)\text{是幺模阵} \tag{28}$$

成立．在该式两边左乘以 $\boldsymbol{P}(s)^{-1}$，右乘以 $\boldsymbol{D}(s)^{-1}\tilde{\boldsymbol{D}}(s)$，得

$$\boldsymbol{P}(s)^{-1}\begin{bmatrix}\tilde{\boldsymbol{D}}(s)\\ \tilde{\boldsymbol{N}}(s)\end{bmatrix}=\begin{bmatrix}\boldsymbol{I}_r\\ \boldsymbol{0}\end{bmatrix}\boldsymbol{Q}(s)\boldsymbol{D}(s)^{-1}\tilde{\boldsymbol{D}}(s) \tag{29}$$

将 $\boldsymbol{P}(s)^{-1}$划分成分块矩阵，设

$$\boldsymbol{P}(s)^{-1}=\begin{array}{c}\begin{bmatrix}\boldsymbol{\Pi}_{11}(s) & \boldsymbol{\Pi}_{12}(s)\\ \underbrace{\boldsymbol{\Pi}_{21}(s)}_{r} & \underbrace{\boldsymbol{\Pi}_{22}(s)}_{m}\end{bmatrix}\end{array}\begin{matrix}\}r\\ \}m\end{matrix} \tag{30}$$

将其代入(29)式，然后对得到的式子再左乘以 $\boldsymbol{Q}(s)^{-1}$，得

$$\boldsymbol{Q}(s)^{-1}\{\boldsymbol{\Pi}_{11}(s)\tilde{\boldsymbol{D}}(s)+\boldsymbol{\Pi}_{12}(s)\tilde{\boldsymbol{N}}(s)\}=\boldsymbol{D}^{-1}(s)\tilde{\boldsymbol{D}}(s) \tag{31}$$

因 $\boldsymbol{P}(s)$，$\boldsymbol{Q}(s)$是幺模阵，则上式左边是多项式矩阵．因此，右边的 $D(s)^{-1}\tilde{\boldsymbol{D}}(s)$ 也是多项式矩阵．

(b)(⇒)设 $\tilde{\boldsymbol{N}}(s)$和 $\tilde{\boldsymbol{D}}(s)$的 gcrd 为 $\tilde{\boldsymbol{R}}(s)$，即令

$$\tilde{\boldsymbol{N}}(s)=\tilde{\boldsymbol{N}}_0(s)\tilde{\boldsymbol{R}}(s),\ \tilde{\boldsymbol{D}}(s)=\tilde{\boldsymbol{D}}_0(s)\tilde{\boldsymbol{R}}(s)$$

根据假定

$$\boldsymbol{N}(s)=\tilde{\boldsymbol{N}}_0(s)\{\tilde{\boldsymbol{R}}(s)\boldsymbol{R}(s)^{-1}\},$$
$$\boldsymbol{D}(s)=\tilde{\boldsymbol{D}}_0(s)\{\tilde{\boldsymbol{R}}(s)\boldsymbol{R}(s)^{-1}\}$$

因 $\boldsymbol{N}(s)$和 $\boldsymbol{D}(s)$互相右素，则 $\{\tilde{\boldsymbol{R}}(s)\boldsymbol{R}(s)^{-1}\}$ 是幺模阵．由假定，$\boldsymbol{R}(s)$ 也是幺模阵，因而 $\tilde{\boldsymbol{R}}(s)$也是幺模阵，即 $\tilde{\boldsymbol{N}}(s)$和 $\tilde{\boldsymbol{D}}(s)$相互右素．

(⇐)由(a)，对于某个多项式矩阵 $\boldsymbol{R}(s)$，$\tilde{\boldsymbol{N}}(s)=\boldsymbol{N}(s)\boldsymbol{R}(s)$，$\tilde{\boldsymbol{D}}(s)=\boldsymbol{D}(s)\boldsymbol{R}(s)$ 成立．因 $\tilde{\boldsymbol{N}}(s)$和 $\tilde{\boldsymbol{D}}(s)$相互右素，则 $\boldsymbol{R}(s)$是幺模阵．

证明完毕

由该定理容易得到下列[系 1]，[系 2]．

[系 1] 设 $\boldsymbol{H}(s)$ 是真正适宜的，并令 $\boldsymbol{H}(s)$ 的既约分解为 $\boldsymbol{H}(s)=\boldsymbol{N}_1(s)\boldsymbol{D}_1(s)^{-1}=\boldsymbol{D}_2(s)^{-1}\boldsymbol{N}_2(s)$，则 $\boldsymbol{N}_1(s)$和 $\boldsymbol{N}_2(s)$，$\boldsymbol{D}_1(s)$ 和 $\boldsymbol{D}_2(s)$ 的不自明(即 1 以外的)不变因子分别一致．

（证明） 利用司密斯典范形，如(25)式，可以得到一个右既约分解 $\boldsymbol{N}_R(s)\boldsymbol{D}_R(s)^{-1}$．完全同样，一个左既约分解 $\boldsymbol{D}_L(s)^{-1}\boldsymbol{N}_L(s)$可给成

$$\boldsymbol{N}_L(s)=\begin{bmatrix}n_1(s) & & & & \\ & \ddots & & \boldsymbol{0} & \\ & & n_R(s) & & \\ & & & 0 & \\ & \boldsymbol{0} & & & \ddots \\ & & & & & 0\end{bmatrix}\boldsymbol{Q}(s),\ \boldsymbol{D}_L(s)=\begin{bmatrix}d_1(s) & & & & \\ & \ddots & & \boldsymbol{0} & \\ & & d_R(s) & & \\ & & & 1 & \\ & \boldsymbol{0} & & & \ddots \\ & & & & & 1\end{bmatrix}\boldsymbol{P}(s)^{-1} \tag{32}$$

其中 $\boldsymbol{N}_L(s)$和 $\boldsymbol{N}_R(s)$ 有同样的不变因子 $n_1(s)$，…，$n_R(s)$，而且 $\boldsymbol{D}_L(s)$ 和 $\boldsymbol{D}_R(s)$ 有除 1 以外的同样的不变因子 $d_1(s)$，…，$d_n(s)$．

根据定理，对于任意右既约分解，$N_1(s)=N_R(s)R_1(s)$，$D_1(s)=D_R(s)R_1(s)$，$R_1(s)$ 是幺模阵．而且，对于任意左既约分解，$N_2(s)=R_2(s)N_L(s)$，$D_2(s)=R_2(s)D_L(s)$，$R_2(s)$ 是幺模阵．因为乘上幺模阵后不变因子不变，则 $N_1(s)$ 的不变因子与 $N_2(s)$ 的不变因子一致．根据上述，后者与 $N_L(s)$ 的不变因子一致，因而和 $N_2(s)$ 的不变因子一致．对于 $D_1(s)$ 和 $D_2(s)$ 也同样．

［系 2］ 设 $H(s)$ 为真正适宜的，$H(s)$ 的既约分解形为 $H(s)=N_1(s)D_1(s)^{-1}=D_2(s)^{-1}N_2(s)$，则其最小实现的维数与 $\det D_1(s)$，$\det D_2(s)$ 的次数一致，$\det D_1(s)$，$\det D_2(s)$ 给出特征多项式的常数倍．

（证明） 只对右分解形证明，对于左分解形也是一样的．首先，因(25)式是一个右既约分解形，根据定理，对于任意右既约分解形，有满足

$$D_R(s)=D_1(s)R_1(s) \tag{33}$$

的幺模阵 $R_1(s)$ 存在．根据前节(18)式和本节(25)，(33)式，

$$\left.\begin{aligned}\text{最小实现的特征多项式}&=d_1(s)d_2(s)\cdots d_R(s)\\&=\det Q(s)\det D_R(s)\\&=\det Q(s)\det R_1(s)\det D_1(s)\end{aligned}\right\} \tag{34}$$

因 $Q(s)$，$R_1(s)$ 均为幺模阵，则"$\det Q(s)\det R_1(s)=$不为 0 的常数"．

证明完毕

系 2 是将起初讲的纯量传递函数的性质推广到矩阵传递函数．

(ii) 和状态变量表示的关系

下面再稍微谈一下状态方程式、输出方程式

$$\dot{x}=Ax+Bu \tag{35}$$

$$y=Cx \tag{36}$$

和传递函数右分解形的关系．这里假定，$x(t)\in R^n$，$u(t)\in R^r$，$y(t)\in R^m$，$\operatorname{rank} C=m$，$\operatorname{rank} B=r$，而且起始条件 $x(0)=0$．首先进行适当的座标变换，假定(1)，(2)式已变成下述形式(例如吕伯格(Luenberger)可控典范形)，即矩阵 A，B 为下列形式

$$A=\begin{bmatrix}A_{11}\cdots A_{1r}\\ \vdots \qquad \vdots\\ A_{r1}\cdots A_{rr}\end{bmatrix},\quad A_{ii}=\begin{bmatrix}0 & 1 & 0 & \cdots & 0\\ \vdots & & \ddots & \ddots & \vdots\\ \vdots & & & \ddots & 0\\ 0 & \cdots & \cdots & 0 & 1\\ a_1^{ii} & \cdots & \cdots & \cdots & a_{\sigma_i}^{ii}\end{bmatrix}=\sigma_i\times\sigma_i,$$

$$A_{ij}=\begin{bmatrix}0 & \cdots & 0\\ \vdots & & \vdots\\ 0 & \cdots & 0\\ a_1^{ij} & \cdots & a_{\sigma_j}^{ij}\end{bmatrix}=\sigma_i\times\sigma_j \tag{37}$$

$$B=\begin{bmatrix}B_1\\ \cdots\cdots\\ \vdots\\ \cdots\cdots\\ B_r\end{bmatrix},\quad B_i=\begin{bmatrix}0 & \cdots & 0\\ \vdots & & \vdots\\ 0 & \cdots & 0\\ b_{i1} & \cdots & b_{ir}\end{bmatrix}=\sigma_i\times r \tag{38}$$

式中 $\sigma_i(i=1, \cdots, r)$ 是称为 $(\boldsymbol{A}, \boldsymbol{B})$ 的克罗内克指数的自然数，当然 $\sum_{i=1}^{r}\sigma_i=n$[1]. 此外假定，$\boldsymbol{x}$，$\boldsymbol{C}$ 也按照 $\boldsymbol{A}$，$\boldsymbol{B}$ 的分割，划分成

$$\boldsymbol{x}=\begin{bmatrix} \boldsymbol{x}^1 \\ \vdots \\ \boldsymbol{x}^r \end{bmatrix}, \quad \boldsymbol{x}^j(t)=\begin{bmatrix} x_1^j(t) \\ \vdots \\ x_{\sigma_j}^j(t) \end{bmatrix} \in R^{\sigma_j} \tag{39}$$

$$\boldsymbol{C}=\begin{bmatrix} \boldsymbol{C}_{11}\cdots\boldsymbol{C}_{1r} \\ \vdots \qquad \vdots \\ \boldsymbol{C}_{m1}\cdots\boldsymbol{C}_{mr} \end{bmatrix}, \quad \boldsymbol{C}_{ij}=[c_1^{ij}\cdots c_{\sigma_j}^{ij}]=1\times\sigma_j, \; j=1, \cdots, r; \; i=1, \cdots, m \tag{40}$$

对(35)，(36)进行拉氏变换，利用上述定义，则不难证明

$$\boldsymbol{y}(s)=\boldsymbol{N}(s)\boldsymbol{D}(s)^{-1}\boldsymbol{u}(s) \tag{41}$$

$$\boldsymbol{N}(s)\triangleq[c_1^{ij}+c_2^{ij}s+\cdots+c_{\sigma_j}^{ij}s^{\sigma_j-1}] \tag{42}$$

$$\boldsymbol{D}(s)\triangleq\boldsymbol{B}_0^{-1}\boldsymbol{D}_0(s) \tag{43}$$

$$\boldsymbol{B}_0\triangleq[b_{ij}] \tag{44}$$

$$\boldsymbol{D}_0(s)\triangleq[-a_1^{ij}-a_2^{ij}s\cdots a_{\sigma_j}^{ij}s^{\sigma_j-1}+\delta_{ij}s^{\sigma_j}]. \tag{45}$$

容易理解，$\boldsymbol{D}(s)$ 第 j 列的次数是 σ_j，$\boldsymbol{N}(s)$ 第 j 列的次数在 (σ_j-1) 以下。

反之，当真正适宜的传递函数 $\boldsymbol{H}(s)$ 以右分解形 $\widetilde{\boldsymbol{N}}(s)\widetilde{\boldsymbol{D}}(s)^{-1}$ 给出时，对应的状态方程式，输出方程式可按下述求出. 因 $\det\widetilde{\boldsymbol{D}}(s)\neq0$，则可利用某个幺模阵 $\boldsymbol{U}(s)$ 使 $\boldsymbol{D}(s)\triangleq\widetilde{\boldsymbol{D}}(s)\boldsymbol{U}(s)$ 变成列适宜的[2]. 注意，若令 $\boldsymbol{N}(s)\triangleq\widetilde{\boldsymbol{N}}(s)\boldsymbol{U}(s)$，根据定理 2，$\boldsymbol{N}(s)\boldsymbol{D}(s)^{-1}$ 也是 $\boldsymbol{H}(s)$ 的右分解形. 这时，若只将 $\boldsymbol{D}(s)$ 各列最高次数元素的系数留下，设所组成的常数矩阵为 $\boldsymbol{P}$，因 $\boldsymbol{D}(s)$ 是列适宜的，则 $\det\boldsymbol{P}\neq0$. 因此，若令 $\boldsymbol{B}_0\triangleq\boldsymbol{P}^{-1}$，$\boldsymbol{D}(s)=\boldsymbol{P}\{\boldsymbol{P}^{-1}\boldsymbol{D}(s)\}=\boldsymbol{B}_0^{-1}\boldsymbol{D}_0(s)$，其中 $\boldsymbol{D}_0(s)\triangleq\boldsymbol{P}^{-1}\boldsymbol{D}(s)$，显然 $\boldsymbol{D}_0(s)$ 和(45)式具有同样形式. 对于由 $\widetilde{\boldsymbol{N}}(s)$，$\widetilde{\boldsymbol{D}}(s)$ 这样求出的 $\boldsymbol{N}(s)$，$\boldsymbol{D}_0(s)$，$\boldsymbol{B}_0$，可以按与上述相反的次序进行.

而且，可观测典范形和左分解形的关系也可以同样推出.

在这里，对于自然数组 $\sigma_1, \cdots, \sigma_r$，定义

$$\boldsymbol{S}(\sigma_1, \cdots, \sigma_r)\triangleq\begin{bmatrix} \begin{matrix}1\\ s\\ \vdots\\ s^{\sigma_1-1}\end{matrix} & & & \boldsymbol{0} \\ & \begin{matrix}1\\ s\\ \vdots\\ s^{\sigma_2-1}\end{matrix} & & \\ & & \ddots & \\ \boldsymbol{0} & & & \begin{matrix}1\\ s\\ \vdots\\ s^{\sigma_r-1}\end{matrix} \end{bmatrix} \tag{46}$$

容易看出，利用该矩阵，(42)、(45)式可写成

1) 参照前面部分。

2) 参照前面部分引理 6。

$$N(s)=CS(\sigma_1,\ \cdots,\ \sigma_r) \tag{47}$$

$$D_0(s)=\mathrm{diag}(s^{\sigma_1},\ \cdots,\ s^{\sigma_r})-A_0S(\sigma_1,\ \cdots,\ \sigma_r) \tag{48}$$

式中

$$A_0 \triangleq \left[\begin{array}{c|c|c} a_1^{11}\cdots a_{\sigma_1}^{11} & \cdots & a_1^{1r}\cdots a_{\sigma_1}^{1r} \\ \hline \vdots & & \vdots \\ \hline a_1^{r1}\cdots a_{\sigma_1}^{r1} & \cdots & a_1^{rr}\cdots a_{\sigma_r}^{rr} \end{array}\right] \tag{49}$$

(iii) 可观测性和可控性

[定理 3] 假定用状态方程式(35)及输出方程式(36)表示的系统的传递函数 $H(s)=C(sI-A)^{-1}B$ 可右分解成

$$H(s)=N(s)D(s)^{-1},\ \delta(\det D(s))=n$$

这时

$N(s)$ 和 $D(s)$ 相互右素⇔(35), (36)式是 $H(s)$ 的最小实现(即 $(A,\ B)$ 可控，且 $(A,\ C)$ 可观测)

对于左分解也有同样关系.

(证明) 根据系 2, 若 $N(s)$ 和 $D(s)$ 相互右素，则 $\delta(\det D(s))$ 是 $H(s)$ 最小实现的次数.

反之,若 $N(s)$ 和 $D(s)$ 不相互右素,表示成

$$N(s)=\tilde{N}(s)R(s),\ D(s)=\tilde{D}(s)R(s),\ \delta(\det R(s))>0$$

则 $H(s)$ 最小实现的次数在 $\delta(\det\tilde{D}(s))$ 以下,若 $\delta(\det\tilde{D}(s))<n$, 则 $(A,\ B,\ C)$ 不是最小实现.

证明完毕

将该定理用于可控性或可观测性之一成立的系统,根据 $N(s)$ 和 $D(s)$ 是相互左素或右素,可以判别其另一个性质是否成立.

[系 3] 对于用吕恩倍格可控典范形表示的系统,

$(A,\ C)$ 可观测⇔由 $(A,\ C)$ 的元素按(42)～(45)式组成的 $N(s)$ 和 $D_0(s)$ 相互右素

(问题 2) 试由吕伯格可观测典范形按(41)～(45)式作出 $H(s)$ 的左分解形，说明可观测典范形可控的条件.

[系 4] 在系统(35), (36)式中:

(a) $(A,\ C)$ 可观测⇔$[sI-A]$ 和 C 相互右素

(b) $(A,\ B)$ 可控⇔$[sI-A]$ 和 B 相互左素

(证明) (a) 代替系统 $(A,\ B,\ C)$, 考虑可控系统 $(A,\ I,\ C)$, 显然 $N(s)=C,\ D(s)=sI-A$ 为传递函数 $C(sI-A)^{-1}I$ 的右分解,而且 $\delta(\det(sI-A))=n$.

(b) 代替系统 (A,B,C), 考虑可观测系统 (A,B,I), 显然 $D(s)=sI-A,\ N(s)=B$ 为传递函数 $I(sI-A)^{-1}B$ 的左分解,按照上述, $\delta(\det(sI-A))=n$.

证明完毕

(注意) 设 $H_0(s)\triangleq(sI-A)^{-1}B$ 的一个右既约分解为 $N(s)D(s)^{-1}$, $D(s)=r\times r$, 根据系 4(b)和定理 2 的系 1 可得出下列关系, “$(A,\ B)$ 可控⇔$(sI-A)^{-1}B$ 是 $H_0(s)$ 的左既约分解⇒$(sI-A)$ 的除 1 以外的不变因子和 $D(s)$ 的除 1 以外的不变因子一致

$\Rightarrow(sI-A)$的除1以外的不变因子的数目在r以下”即“为使系统在r个输入时可控，则$(sI-A)$除1以外的不变因子的数目必须在r以下”，当然对于可观测性也是一样。

(iv) 状态反馈

下面我们来研究一下对系统(35)，(36)式进行线性状态反馈的效果。在这里同样假定$\boldsymbol{A}$, $\boldsymbol{B}$选成(37)，(38)式的形式。设利用(44)式的$\boldsymbol{B}_0$，反馈可以表示成

$$\boldsymbol{u}=\boldsymbol{B}_0^{-1}\boldsymbol{K}\boldsymbol{x}+\boldsymbol{G}\boldsymbol{v},\ \det\boldsymbol{G}\neq 0 \tag{50}$$

再分割$\boldsymbol{K}$，令

$$\boldsymbol{K}=\begin{bmatrix}\boldsymbol{k}_{11}\cdots\boldsymbol{k}_{1r}\\ \vdots \qquad \vdots\\ \boldsymbol{k}_{r1}\cdots\boldsymbol{k}_{rr}\end{bmatrix},\ \boldsymbol{k}_{ij}=[k_1^{ij},\ \cdots,\ k_{\sigma_j}^{ij}]=1\times\sigma_j \tag{51}$$

将(50)，(51)式代入(35)，(36)式，得到的系统为

$$\dot{\boldsymbol{x}}=(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{B}_0^{-1}\boldsymbol{K})\boldsymbol{x}+\boldsymbol{B}\boldsymbol{G}\boldsymbol{v} \tag{52}$$

$$\boldsymbol{y}=\boldsymbol{C}\boldsymbol{x} \tag{53}$$

容易看出，矩阵$(\boldsymbol{A}+\boldsymbol{B}\boldsymbol{B}_0^{-1}\boldsymbol{K})$和$\boldsymbol{A}$具有同样的结构，只不过是将$a_l^{ij}$换成$(a_l^{ij}+k_l^{ij})$ $(l=1,\cdots,\sigma_j;i,j=1,\cdots,r)$。因此，根据(52)，(53)式，若按着和前面同样的要领求右分解形，则得

$$\boldsymbol{C}(s\boldsymbol{I}-\boldsymbol{A}-\boldsymbol{B}\boldsymbol{B}_0^{-1}\boldsymbol{K})^{-1}\boldsymbol{B}\boldsymbol{G}=\boldsymbol{N}(s)\boldsymbol{D}_K(s)^{-1} \tag{54}$$

$$\boldsymbol{D}_K(s)=\boldsymbol{G}^{-1}\boldsymbol{B}_0^{-1}\boldsymbol{D}_{K0}(s);$$

$$\boldsymbol{D}_{K0}(s)=[-(a_1^{ij}+k_1^{ij})-(a_2^{ij}+k_2^{ij})s-\cdots-(a_{\sigma_j}^{ij}+k_{\sigma_j}^{ij})s^{\sigma_j-1}+\delta_{ij}s^{\sigma_j}] \tag{55}$$

$\boldsymbol{N}(s)$可用(42)式给出。因此，若定义

$$\boldsymbol{K}(s)\triangleq[k_1^{ij}+k_2^{ij}s+\cdots+k_{\sigma_j}^{ij}s^{\sigma_j-1}]=\boldsymbol{K}\boldsymbol{S}(\sigma_1,\ \cdots,\ \sigma_r) \tag{56}$$

利用(45)式的$\boldsymbol{D}_0(s)$和该式的$\boldsymbol{K}(s)$，则$\boldsymbol{D}_{K0}(s)$可写成

$$\boldsymbol{D}_{K0}(s)=\boldsymbol{D}_0(s)-\boldsymbol{K}(s) \tag{57}$$

因此

$$\boldsymbol{N}(s)\boldsymbol{D}_K(s)^{-1}=\boldsymbol{N}(s)[\boldsymbol{G}^{-1}\boldsymbol{B}_0^{-1}\{\boldsymbol{D}_0(s)-\boldsymbol{K}(s)\}]^{-1}. \tag{58}$$

和纯量传递函数的情况一样，状态反馈的效果仅出现在“分母”$\boldsymbol{D}(s)$，而“分子”$\boldsymbol{N}(s)$完全不受影响。

在这里，作为反馈矩阵，我们这样来选择$\hat{\boldsymbol{K}}\hat{\boldsymbol{G}}$：

$$\hat{k}_l^{ij}=-a_l^{ij}\quad l=1,\ \cdots,\ \sigma_j;\ i,\ j=1,\ \cdots,\ r \tag{59}$$

$$\hat{\boldsymbol{G}}=\boldsymbol{I}_r, \tag{60}$$

研究一下这种特殊情况，则

$$\boldsymbol{D}_{K0}=\boldsymbol{\Lambda}(s),$$

而

$$\boldsymbol{\Lambda}(s)\triangleq\operatorname{diag}(s^{\sigma_1},\ \cdots,\ s^{\sigma_r}) \tag{61}$$

在这里我们利用前部分的定理5，即“当给出(61)式的$\boldsymbol{\Lambda}(s)$和满足前部分的(29)，(30)式的首一多项式$\psi_1(s)$，$\psi_2(s)$，$\cdots$，$\psi_r(s)$时，在$\boldsymbol{\Lambda}(s)$上加上适当的$\tilde{\boldsymbol{K}}(s)$，则$\boldsymbol{D}_{K0}(s)=\boldsymbol{\Lambda}(s)+\tilde{\boldsymbol{K}}(s)$是列(行)适宜的，以$\psi_1(s)$，$\psi_2(s)$，$\cdots$，$\psi_r(s)$为不变因子”

利用该结果得下列定理。

[定理4] 对于给定的系统$\boldsymbol{N}(s)\boldsymbol{D}(s)^{-1}$，进行状态反馈$\boldsymbol{K}(s)\triangleq\{\hat{\boldsymbol{K}}(s)+\tilde{\boldsymbol{K}}(s)\}$，

$G\triangleq I_r$, 在所得到的闭环系统 $N(s)D_K(s)^{-1}$ 中,可以使"分母" $D_K(s)=B_0^{-1}\{D_0(s)-\hat{K}(s)-\tilde{K}(s)\}=B_0^{-1}\{\Lambda(s)-\tilde{K}(s)\}=B_0^{-1}D_{K0}(s)$ 的不变因子和满足上述条件的任意首一多项式一致.

特别是,在 $\psi_1(s)=\psi_2(s)=\cdots=\psi_{r-1}(s)=1$, $\psi_r(s)=s^2+a_{n-1}s^{n-1}+\cdots+\alpha_1 s+\alpha_0$ 和仅指定特征多项式的情况下(极点配置问题),对于 $\tilde{K}$ 可选择[1)]

$$\tilde{K}(s)=\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \\ 0 & \cdots & \cdots & 0 & 1 \\ k_1(s) & \cdots & & k_{r-1}(s) & k_r(s) \end{bmatrix}$$

$$\begin{aligned} k_1(s) &= \alpha_0+\alpha_1(s)+\cdots+\alpha_{\sigma_1-1}s^{\sigma_1-1} \\ k_2(s) &= \alpha_{\sigma_1}+\alpha_{\sigma_1+1}s+\cdots+\alpha_{\sigma_1+\sigma_2-1}s^{\sigma_2-1} \\ &\vdots \\ k_r(s) &= \alpha_{\sigma_1+\cdots+\sigma r-1}+\alpha_{\sigma_1+\cdots+\sigma_{r-1}+1}s+\cdots+\alpha_n s^{\sigma_r-1} \end{aligned}$$

如系 3 后面的注意中所述,若 (A, B) 可控,则 $(sI-A)$ 的除 1 以外的不变因子与 $D(s)$ 的除 1 以外的不变因子一致. 若 (A, B) 可控,则对于任意 F, $(A+BF, B)$ 仍然可控. 考虑到以上各点,由定理 4 可得下列结果.

[系 5] 若 (A, B) 可控,则可以选择适当的 F, 使 $(sI-A-BF)$ 的不变因子和满足上述条件的任意首一多项式一致.

参 考 文 献

[1] L. A. Zadeh & C. A. Desoer: Linear System Theory; McGraw Hill (1963).

[2] E. Polak & E. Wong: Notes for A First Course on Linear Systems, Van Nostrand Reinhold (1970).

[3] 古田: システムの実現問題; 計測と制御, Vol. 13, No. 6, pp. 497~508(昭和 49).

[4] M. Heymann & J. Thorpe: Transfer Equivalence of Linear Dynamical Systems; SIAM J. Control, Vol. 8, No. 1, pp. 19~40, Feb. (1960).

[5] C. T. Chen: Representations of Linear Time-Invariant Composite Systems; IEEE Trans. Automatic Control, Vol. AC-13, No. 3, pp. 277~283 (1968).

[6] D. G. Luenberger: Canonical Forms for Linear Multi-Variable Systems; ibid., Vol. AC-12, pp. 290~293 (1967).

(参照 A-I-12)

附录(引理 3 的证明)

(i) 一般,当设 $d_1(s)$ 为 A 的最小多项式时,则

$$s^k d_1(s)(sI-A)^{-1}=A^k d_1(s)(sI-A)^{-1}\bmod d_1(s),\ k=1, 2, 3, \cdots$$

成立. 这是因为,该式对于 $k=1$ 是自明的;若设 $k=l$ 时上式成立,则

$$\begin{aligned} s^{l+1}d_1(s)(sI-A)^{-1}-A^{l+1}d_1(s)(sI-A)^{-1} &= s^l(sI-A)d_1(s)(sI-A)^{-1} \\ &= s^l I d_1(s)=0 \quad \bmod d_1(s), \end{aligned}$$

故对于 $k=l+1$ 也成立.

1) 参照 A-I-13 中(39)式.

由此结果得

$$d_1(s)\boldsymbol{C}(s\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}=d_1(s)\,(\Sigma \boldsymbol{P}_i\tilde{\boldsymbol{C}}\tilde{\boldsymbol{A}}^i)\,(s\boldsymbol{I}-\tilde{\boldsymbol{A}})^{-1}(\Sigma\tilde{\boldsymbol{A}}^i\tilde{\boldsymbol{B}}\boldsymbol{Q}_i)$$
$$=d_1(s)\boldsymbol{P}(s)\tilde{\boldsymbol{C}}(s\boldsymbol{I}-\tilde{\boldsymbol{A}})^{-1}\tilde{\boldsymbol{B}}\boldsymbol{Q}(s)\quad \operatorname{mod} d_1(s)\qquad (\mathrm{A}.1)$$

(ii) 首先证明 $S=$可控且可观测$\Rightarrow\hat{S}$ 可控且可观测.

对于某个向量 $\boldsymbol{x}\in R^n$, 令

$$\boldsymbol{x}[\tilde{\boldsymbol{B}}\ \tilde{\boldsymbol{A}}\tilde{\boldsymbol{B}},\ \cdots,\ \tilde{\boldsymbol{A}}^{n-1}\tilde{\boldsymbol{B}}]=\boldsymbol{0}(\tilde{\boldsymbol{A}}\in R^{n\times n},\ \tilde{\boldsymbol{B}}\in R^{n\times r})$$

根据凯莱-哈密顿定理, $\boldsymbol{x}\tilde{\boldsymbol{A}}^k\tilde{\boldsymbol{B}}=\boldsymbol{0}$, $k=1,\ 2,\ \cdots$. 这意味着, 再由(16)式, $\boldsymbol{x}\boldsymbol{A}^k\boldsymbol{B}=\boldsymbol{0}$, $k=1,\ 2,\ \cdots$, 即证明 $\boldsymbol{x}[\boldsymbol{B},\ \boldsymbol{A}\boldsymbol{B},\ \cdots,\ \boldsymbol{A}^{n-1}\boldsymbol{B}]=\boldsymbol{0}$,

$\operatorname{rank}[\tilde{\boldsymbol{B}},\tilde{\boldsymbol{A}}\tilde{\boldsymbol{B}},\cdots,\ \tilde{\boldsymbol{A}}^{n-1}\tilde{\boldsymbol{B}}]\geqslant\operatorname{rank}[\boldsymbol{B},\boldsymbol{A}\boldsymbol{B},\cdots,\ \boldsymbol{A}^{n-1}\boldsymbol{B}]$或 $S=$可控$\Rightarrow\tilde{S}=$可控. 同样证明 $S=$可观测$\Rightarrow\tilde{S}=$可观测.

下面证明其逆. 根据(A.1),

$$d_1(s)\tilde{\boldsymbol{C}}(s\boldsymbol{I}-\tilde{\boldsymbol{A}})^{-1}\tilde{\boldsymbol{B}}=d_1(s)\boldsymbol{P}(s)^{-1}\boldsymbol{C}(s\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}\boldsymbol{Q}(s)^{-1}\quad \operatorname{mod} d_1(s)$$

成立. 因 $\boldsymbol{P}(s)$, $\boldsymbol{Q}(s)$是幺模阵, 则可表示成 $\boldsymbol{P}(s)^{-1}=\Sigma\tilde{\boldsymbol{P}}_is^i$, $\boldsymbol{Q}(s)^{-1}=\Sigma\tilde{\boldsymbol{Q}}_is^i$. 将其代入上式,得

$$d_1(s)\tilde{\boldsymbol{C}}(s\boldsymbol{I}-\tilde{\boldsymbol{A}})^{-1}\tilde{\boldsymbol{B}}=d_1(s)\,(\Sigma\tilde{\boldsymbol{P}}_is^i)\boldsymbol{C}(s\boldsymbol{I}-\boldsymbol{A})^{-1}\boldsymbol{B}(\Sigma\tilde{\boldsymbol{Q}}_is^i)\quad \operatorname{mod} d_1(s)$$
$$=d_1(s)\,(\Sigma\tilde{\boldsymbol{P}}_i\boldsymbol{C}\boldsymbol{A}^i)\,(s\boldsymbol{I}-\boldsymbol{A})^{-1}(\Sigma\boldsymbol{A}^i\boldsymbol{B}\tilde{\boldsymbol{Q}}_i)\quad \operatorname{mod} d_1(s)$$

这意味着

$$\tilde{\boldsymbol{C}}(s\boldsymbol{I}-\tilde{\boldsymbol{A}})^{-1}\tilde{\boldsymbol{B}}=(\Sigma\tilde{\boldsymbol{P}}_i\boldsymbol{C}\boldsymbol{A}^i)\,(s\boldsymbol{I}-\boldsymbol{A})^{-1}(\Sigma\boldsymbol{A}^i\boldsymbol{B}\tilde{\boldsymbol{Q}}_i)\qquad (\mathrm{A}.2)$$

现假定 $\tilde{S}=(\tilde{\boldsymbol{A}},\tilde{\boldsymbol{B}},\tilde{\boldsymbol{C}})$可控且可观测, 根据(A.2), $(\boldsymbol{A},\Sigma\boldsymbol{A}^i\boldsymbol{B}\tilde{\boldsymbol{Q}}_i,\Sigma\tilde{\boldsymbol{P}}_i\boldsymbol{C}\boldsymbol{A}^i)$也是 $\tilde{S}$ 的传递函数的最小实现.因此,对于某个正则矩阵 $\boldsymbol{T}\in\boldsymbol{R}^{n\times n}$, $\tilde{\boldsymbol{B}}=\boldsymbol{T}^{-1}\Sigma\boldsymbol{A}^i\boldsymbol{B}\tilde{\boldsymbol{Q}}_i$, $\tilde{\boldsymbol{C}}=(\Sigma\tilde{\boldsymbol{P}}_i\boldsymbol{C}\boldsymbol{A}^i)\boldsymbol{T}$. 和前面一样,这个关系意味着,$\tilde{S}=$可控且可观测$\Rightarrow S=$可控且可观测.